

STIC-Biotech/ChemLib

165565

mg

From: Whiteman, Brian  
Sent: Tuesday, September 13, 2005 1:42 PM  
To: STIC-Biotech/ChemLib  
Subject: seq search

RECEIVED  
SEP 13 2005  
STIC/BIOTECH DIVISION  
(STIC)

09/380,203 La Monte et al. 2/26/98

search seq id no: 1 against DNA

1) the commercial databases, and the issued and published US application databases

search seq id no: 2 against dna and amino acid

1) the commercial databases, and the issued and published US application databases

Thank you,

Brian Whiteman  
Remsen, 2D14  
mail box 2C18  
Patent Examiner - Art Unit 1635  
United States Patent and Trademark Office  
(571) 272-0764

9/13/05  
1-NA - QSH  
1-AA - QAP  
1-AA - reverse to  
NA

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 15, 2005, 20:33:38 ; Search time 5428 Seconds  
(without alignments)  
3347.588 Million cell updates/sec

Title: US-09-380-203-2

Perfect score: 2034

Sequence: 1 MEFSLLPRLECNCAISAHK.....FIRGGVSPVLSGWSQTPDLR 375

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame\_p2n.model -DEV=xlh  
-Q/cgn2\_1/USPTO.spool/US09380203/runat\_13092005\_170810\_11670/app.query.fasta\_1.519  
-DB=GenEmbl -QFMT=fasap -SUFFIX=rg -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09380203 @CNG 1.1 4200 @runat\_13092005\_170810\_11670 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2034	100.0	1442	6	AR051550 Sequence
2	2034	100.0	1442	6	AR072690 Sequence
3	2034	100.0	1442	6	AR073235 Sequence
4	2034	100.0	1442	6	AX594290 Sequence

5	2034	100.0	1442	6	AX599994	Sequence
6	2034	100.0	1442	6	AX709024	Sequence
7	2034	100.0	1442	6	AX709356	Sequence
8	2034	100.0	1442	6	AX770807	Sequence
9	2034	100.0	1442	6	AX774635	Sequence
10	2034	100.0	1442	6	BD070526	Transgeni
11	2034	100.0	1442	9	AF010144	Homo sapi
12	1799	88.4	124001	9	AF010144	Homo sapi
13	1787	87.9	1418	6	HS886K2	Transgeni
14	1520	74.7	1381	6	BD070528	Transgeni
15	1520	74.7	1381	6	AR051479	Sequence
16	1520	74.7	1381	6	AR072619	Sequence
17	1520	74.7	1381	6	AR073164	Sequence
18	982.5	48.3	160597	9	BD070527	Transgeni
19	977.5	48.1	62195	9	AL590666	Human DNA
20	968	47.6	220384	2	AL109952	Human DNA
21	967.5	47.6	108921	9	AC078926	Human sapi
22	967.5	47.6	189729	2	AC004263	Homo sapi
23	954.5	46.9	209252	9	AC073930	Homo sapi
24	949	46.7	72826	9	AC097637	Homo sapi
25	937	45.1	174617	2	AL391539	Human DNA
26	925.5	45.5	195558	9	AC023671	Homo sapi
27	925	45.5	108879	9	AC007448	Homo sapi
28	925	45.5	240058	9	AC116491	Homo sapi
29	922	45.3	115932	9	AC110921	Homo sapi
30	921.5	45.3	135975	9	AC011446	Homo sapi
31	914	44.9	38875	9	AC105935	Homo sapi
32	913.5	44.9	54277	9	AC092101	Homo sapi
33	913.5	44.9	150562	2	AC133473	Homo sapi
34	913.5	44.9	208273	2	AC069420	Homo sapi
35	910.5	44.8	44936	9	AC134942	Homo sapi
36	909	44.7	175827	9	AC007774	Homo sapi
37	908	44.6	163424	2	AL358472	Human DNA
38	907.5	44.6	189317	9	AC099330	Homo sapi
39	907	44.6	65609	9	HS0190720	Human DNA
40	903.5	44.4	172601	2	AC025858	Homo sapi
41	902	44.3	83840	9	AC021455	Homo sapi
42	901.5	44.3	188453	9	HS0148H17	Human DNA
43	901	44.3	33362	6	AC025521	Homo sapi
44	901	44.3	92797	9	C0869844	Sequence
45	900.5	44.3	173316	2	AF038458	Homo sapi
					AC025983	Homo sapi

#### ALIGNMENTS

RESULT 1	AR051550	Sequence 120 from patent US 5830670.	1442 bp	DNA	linear	PAT 29-SEP-1999
LOCUS	AR051550	Sequence 120 from patent US 5830670.	1442 bp	DNA	linear	PAT 29-SEP-1999
DEFINITION	AR051550	Sequence 120 from patent US 5830670.	1442 bp	DNA	linear	PAT 29-SEP-1999
ACCESSION	AR051550	Sequence 120 from patent US 5830670.	1442 bp	DNA	linear	PAT 29-SEP-1999
VERSION	AR051550.1	GI:5974914	1442 bp	DNA	linear	PAT 29-SEP-1999
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 1442)					
AUTHORS	de la Monte, S. and Wande, J.R.					
TITLE	Neural thread protein gene expression and detection of Alzheimer's disease					
JOURNAL	Patent: US 5830670-A 120 03-NOV-1998;					
FEATURES	Location/Qualifiers					
source	1..1442					
	/organism="unknown"					
	/mol_type="unassigned DNA"					
ORIGIN						
Alignment Scores:						
Pred. No.:	4,69e-142	Length:	1442			
Score:	2034.00	Matches:	375			
Percent Similarity:	100.00%	Conservative:	0			
Best local Similarity:	100.00%	Mismatches:	0			
Query Match:	100.00%	Indels:	0			
DB:	6	Gaps:	0			

US-09-380-203-2 (1-375) x AR051550 (1-1442)		
QY	1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg	20
Db	15 ATGGAGTTTCGCTCTGTTGTCGCCAGGCTGGAGTGCAATGGCGCAATCTCAGCTCACCGC	74
QY	21 AenLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle	40
Db	75 AACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCGCTCAGCTCCCGCTCAGCTGGGATT	134
QY	41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyPhePheLeuValGluMetGluPhe	60
Db	135 ACAGGCATGTGCACCGCTCGGCTAAATTTTGTATTTTATTTAGTAGAGATGGAGTTT	194
QY	61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspProSerValSerAla	80
Db	195 CTCCATGTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCCCTCCGCTCGGCC	254
QY	81 SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys	100
Db	255 TCCCAAGTGTAGATACAGGACTGGCCACCATGCGCGCTCTGCTGGCTAAATTTTGT	314
QY	101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr	120
Db	315 GGTAGAAACAGGGTTTCATCTGATGTGCCAAGCTGGTCTCTCTGAGCTCAAGCAGTCCAC	374
QY	121 CysLeuSerLeuProLysCysTrpAspTyrArgAlaAlaValProGlyLeuPheIle	140
Db	375 TGCCCTCAGCTCCCAAGTGTGGGATTTACAGCGGTGACGGCTGCGCTGGCCCTTTTATT	434
QY	141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp	160
Db	435 TTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAGTGCAGTGGTGTGAT	494
QY	161 HisSerSerLeuGlnProSerThrProGluIleLysHisProProAlaSerAlaSerGln	180
Db	495 CACAGCTCACTGCAGCTTCAACTCTCTGAGATCAAGCATCTCTCTCTCAGCTCAGCTCCCAA	554
US-09-380-203-2 (1-375) x AR072690 (1-1442)		
QY	1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg	20
Db	15 ATGGAGTTTCGCTCTGTTGTCGCCAGGCTGGAGTGCAATGGCGCAATCTCAGCTCACCGC	74
QY	21 AenLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle	40
Db	75 AACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCGCTCAGCTCCCGCTCAGCTGGGATT	134
QY	41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyPhePheLeuValGluMetGluPhe	60
Db	135 ACAGGCATGTGCACCGCTCGGCTAAATTTTGTATTTTATTTAGTAGAGATGGAGTTT	194
QY	61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspProSerValSerAla	80
Db	195 CTCCATGTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCCCTCCGCTCGGCC	254
QY	81 SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys	100
Db	255 TCCCAAGTGTAGATACAGGACTGGCCACCATGCGCGCTCTGCTGGCTAAATTTTGT	314
QY	101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr	120
Db	315 GGTAGAAACAGGGTTTCATCTGATGTGCCAAGCTGGTCTCTCTGAGCTCAAGCAGTCCAC	374
QY	121 CysLeuSerLeuProLysCysTrpAspTyrArgAlaAlaValProGlyLeuPheIle	140
Db	375 TGCCCTCAGCTCCCAAGTGTGGGATTTACAGCGGTGACGGCTGCGCTGGCTTTTATT	434
QY	141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp	160
Db	435 TTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAGTGCAGTGGTGTGAT	494
QY	161 HisSerSerLeuGlnProSerThrProGluIleLysHisProProAlaSerAlaSerGln	180
Db	495 CACAGCTCACTGCAGCTTCAACTCTCTGAGATCAAGCATCTCTCTCTCAGCTCAGCTCCCAA	554
QY	181 ValAlaGlyThrLysAspMetHisIstYrThrTrpLeuIlePheIlePheIlePheAsn	200
Db	555 GTAGCTGGGACCAAGACATGACCACTACACTGCTGGCTAAATTTTATTTTATTTTAAAT	614
QY	201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly	220
Db	615 TTTTGTAGACAGCTCAACTCTGTACCCAGGCTGGAGTGGAGTGGCGCAATCTTGGC	674
QY	221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer	240
Db	675 TCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCCAGGCTCTCTGAGTAGC	734
QY	241 TrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeuValGluMetGly	260
Db	735 TGGGACTACAGGCGGCCACACGCTAGCTAAATTTTGTATTTTATTTAGTAGAGATGGGG	794
QY	261 PheThrMetPheAlaArgLeuIleLeuSerGlyProCysAspLeuProAlaSerAla	280
Db	795 TTACCATGTGTGCCAGGTGTGATCTGTGACCTTGTGATCTGCTCTGCTCGGCC	854
QY	281 SerGlnSerAlaGlyIleThrGlyValSerHisAlaArgLeuIlePheAsnPheCys	300
Db	855 TCCCAAGTGTGGGATTACAGGCTGAGCCACCGCCCGCTTATTTTAAATTTTGT	914
QY	301 LeuPheGluMetGlnSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly	320
Db	915 TTGTTTGAATGGAATCTCACTCTGTACCCAGGCTGGAGTGCATGGGCAAAATCTCGGC	974
QY	321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer	340
Db	975 TCATCGAACCTCTGCTCCCGGCTCAAGCGCATCTCTCTCTCAGCTTCCCAAGCAGC	1034
QY	341 TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGlyGly	360
US-09-380-203-2 (1-375) x AR051550 (1-1442)		
QY	1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg	20
Db	15 ATGGAGTTTCGCTCTGTTGTCGCCAGGCTGGAGTGCAATGGCGCAATCTCAGCTCACCGC	74
QY	21 AenLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle	40
Db	75 AACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCGCTCAGCTCCCGCTCAGCTGGGATT	134
QY	41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyPhePheLeuValGluMetGluPhe	60
Db	135 ACAGGCATGTGCACCGCTCGGCTAAATTTTGTATTTTATTTAGTAGAGATGGAGTTT	194
QY	61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspProSerValSerAla	80
Db	195 CTCCATGTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCCCTCCGCTCGGCC	254
QY	81 SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys	100
Db	255 TCCCAAGTGTAGATACAGGACTGGCCACCATGCGCGCTCTGCTGGCTAAATTTTGT	314
QY	101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr	120
Db	315 GGTAGAAACAGGGTTTCATCTGATGTGCCAAGCTGGTCTCTCTGAGCTCAAGCAGTCCAC	374
QY	121 CysLeuSerLeuProLysCysTrpAspTyrArgAlaAlaValProGlyLeuPheIle	140
Db	375 TGCCCTCAGCTCCCAAGTGTGGGATTTACAGCGGTGACGGCTGCGCTGGCTTTTATT	434
QY	141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp	160
Db	435 TTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAGTGCAGTGGTGTGAT	494
QY	161 HisSerSerLeuGlnProSerThrProGluIleLysHisProProAlaSerAlaSerGln	180
Db	495 CACAGCTCACTGCAGCTTCAACTCTCTGAGATCAAGCATCTCTCTCTCAGCTCAGCTCCCAA	554
QY	181 ValAlaGlyThrLysAspMetHisIstYrThrTrpLeuIlePheIlePheIlePheAsn	200
Db	555 GTAGCTGGGACCAAGACATGACCACTACACTGCTGGCTAAATTTTATTTTATTTTAAAT	614
QY	201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly	220
Db	615 TTTTGTAGACAGCTCAACTCTGTACCCAGGCTGGAGTGGAGTGGCGCAATCTTGGC	674
QY	221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer	240
Db	675 TCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCCAGGCTCTCTGAGTAGC	734
QY	241 TrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeuValGluMetGly	260
Db	735 TGGGACTACAGGCGGCCACACGCTAGCTAAATTTTGTATTTTATTTAGTAGAGATGGGG	794
QY	261 PheThrMetPheAlaArgLeuIleLeuSerGlyProCysAspLeuProAlaSerAla	280
Db	795 TTACCATGTGTGCCAGGTGTGATCTGTGACCTTGTGATCTGCTCTGCTCGGCC	854
QY	281 SerGlnSerAlaGlyIleThrGlyValSerHisAlaArgLeuIlePheAsnPheCys	300
Db	855 TCCCAAGTGTGGGATTACAGGCTGAGCCACCGCCCGCTTATTTTAAATTTTGT	914
QY	301 LeuPheGluMetGlnSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly	320
Db	915 TTGTTTGAATGGAATCTCACTCTGTACCCAGGCTGGAGTGCATGGGCAAAATCTCGGC	974
QY	321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer	340
Db	975 TCATCGAACCTCTGCTCCCGGCTCAAGCGCATCTCTCTCTCAGCTTCCCAAGCAGC	1034
QY	341 TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGlyGly	360

Db	1035 TGGGATTAGGGCACCTGCCACCACCCCGCTAAATTTTGTATTTTTCATTAGAGCGGG	1094
QY	361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg	375
Db	1095 GTTTCACCATATTTGTCAGGCTGGTCTCAAACTCCTGACCTCAGG	1139
RESULT 2		
AR072690	LOCUS	AR072690 1442 bp DNA linear PAT 28-AUG-2000
DEFINITION	Sequence 120 from patent US 5948634.	
ACCESSION	AR072690	
VERSION	AR072690.1	GI:9999454
KEYWORDS	.	
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	1 (bases 1 to 1442)	
AUTHORS	de la Monte,S. and Wands,J.R.	
TITLE	Neural thread protein gene expression and detection of alzheimer's	
JOURNAL	disease	
FEATURES	Patent: US 5948634-A 120 07-SEP-1999;	
source	Location/Qualifiers	
	1..1442	
	/organism="unknown"	
ORIGIN	/mol_type="unassigned DNA"	
Alignment Scores:		
Pred. No.:	4..69e-142	Length: 1442
Score:	2034.00	Matches: 375
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	6	Gaps: 0
US-09-380-203-2 (1-375) x AR072690 (1-1442)		
QY	1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg	20
Db	15 ATGGAGTTTCGCTCTGTTGTCGCCAGGCTGGAGTGCAATGGCGCAATCTCAGCTCACCGC	74
QY	21 AenLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle	40
Db	75 AACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCGCTCAGCTCCCGCTCAGCTGGGATT	134
QY	41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyPhePheLeuValGluMetGluPhe	60
Db	135 ACAGGCATGTGCACCGCTCGGCTAAATTTTGTATTTTATTTAGTAGAGATGGAGTTT	194
QY	61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspProSerValSerAla	80
Db	195 CTCCATGTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCCCTCCGCTCGGCC	254
QY	81 SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys	100
Db	255 TCCCAAGTGTAGATACAGGACTGGCCACCATGCGCGCTCTGCTGGCTAAATTTTGT	314
QY	101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr	120
Db	315 GGTAGAAACAGGGTTTCATCTGATGTGCCAAGCTGGTCTCTCTGAGCTCAAGCAGTCCAC	374
QY	121 CysLeuSerLeuProLysCysTrpAspTyrArgAlaAlaValProGlyLeuPheIle	140
Db	375 TGCCCTCAGCTCCCAAGTGTGGGATTTACAGCGGTGACGGCTGCGCTGGCCCTTTTATT	434
QY	141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp	160
Db	435 TTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAGTGCAGTGGTGTGAT	494
QY	161 HisSerSerLeuGlnProSerThrProGluIleLysHisProProAlaSerAlaSerGln	180
Db	495 CACAGCTCACTGCAGCTTCAACTCTCTGAGATCAAGCATCTCTCTCTCAGCTCAGCTCCCAA	554



```

QY 181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIlePheAsn 200
Db |||||
555 GTAGCTGGGACCAAGACATGACCACTACACTGGCTTAATTTTATTTTATTTTAT 614
QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db |||||
615 TTTTGTGAGACAGAGTCTCAACTCTGTCAACCCAGGCTGGAGTGCAGTGGCGCAATCTTGGC 674
QY 221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
Db |||||
675 TCATGCAACCTCTGCTCCGGGTTCAAGTTATTTCTCTGCCAGGCTCTCTGAGTAGC 734
QY 241 TrpAspTyrArgArgProArgLeuAlaAsnPhePheValPheLeuValGluMetGly 260
Db |||||
735 TGGGACTACAGGCGCCACACAGCCTAGCTAATTTTGTATTTTGTATTTTGTAGTAGAGTGGG 794
QY 261 PheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeuProAlaSerAla 280
Db |||||
795 TTCACCATGTTGCCAGAGTGTATCTTGATCTCTGGACCTTGTGATCTGCTGCTGCTGCGGC 854
QY 281 SerGlnSerAlaGlyIleThrGlyValSerHisHisAlaArgLeuIlePheAsnPheCys 300
Db |||||
855 TCCCAAGTGTGGGATTAACAGCGTGAAGCCACCGCCGGCTTAATTTTATTTTGT 914
QY 301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320
Db |||||
915 TTGTTTGAATGGAATCTCACTCTGTATCCAGGCTGGAGTGAATGGCCAAATCTCGGC 974
QY 321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340
Db |||||
975 TCACGTCAACCTCTGCTCCGGGCTCAAGCGATTTCTCTGTCTAGCTTCCCAAGCAGC 1034
QY 341 TrpAspTyrGlyHisLeuProHisProAlaAsnPheCysIlePheIleArgGlyGly 360
Db |||||
1035 TGGGATTACGGGACCTGCCACCAACCCCGCTAATTTTGTATTTTGTATTTTGTATTTTGTATTTAGAGCGGG 1094
QY 361 ValSerProTyrLeuSerGlyTrpSerGlnTrpProAspLeuArg 375
Db |||||
1095 GTTTCACCATATTTGTGAGGCTGGTCTCAACTCTGACCTCAGG 1139

RESULT 3
AR073235
LOCUS AR073235 1442 bp DNA linear PAT 28-AUG-2000
DEFINITION Sequence 120 from patent US 5948888.
ACCESSION AR073235
VERSION AR073235.1 GI:9999998
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1442)
AUTHORS de la Monte, S. and Wands, J.R.
TITLE Neutral thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5948888-A 120 07-SEP-1999;
FEATURES Location/Qualifiers
Source 1..1442
/mol_type="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 4,696-142 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-380-203-2 (1-375) x AR073235 (1-1442)
QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyValAlaIleSerAlaHisArg 20
Db |||||

```

```

15 ATGAGGATTTTCGCTCTTGTGTTGCCAGGCTGGAGTGCAATGGCGCAATCTCAGCTCACCGC 74
QY |||||
21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
Db |||||
75 AACCTCCGCTCCCGGTTCAAGCGATTTCTCTGCTCAGCTCCCGCAGTAGCTGGATT 134
QY |||||
41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60
Db |||||
135 ACAGGCATGTGCACCCAGCTCGGCTAATTTTGTATTTTGTATTTTGTAGTAGAGTGGATT 194
QY |||||
61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAla 80
Db |||||
195 CTCATGTTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCCCTCCGCTCGGCC 254
QY |||||
81 SerGlnSerAlaArgTyrArgThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCys 100
Db |||||
255 TCCCAAGTGTAGATACAGGACTGGCCACCATGCCGCTCTGCTGCTGCTAATTTTGT 314
QY |||||
101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120
Db |||||
315 GGTGAAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTCTGAGCTCAAGCAGTCCACC 374
QY |||||
121 CysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheIle 140
Db |||||
375 TGCCTCAGCTCCCAAGTGTGGGATTACAGCGCTGCAGCGTGCCTGGCCTTTTATTT 434
QY |||||
141 LeuPheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
Db |||||
435 TTATTTTATTTTAAAGACACAGGTGCTCCACTCTTACCCAGGATGAAGTGCAGTGGTGAT 494
QY |||||
161 HisSerSerLeuGlnProSerThrProGluIleLysHisProProAlaSerAlaSerGln 180
Db |||||
495 CACAGCTCACTGCAGCTTCAACTCTCTGAGATCAAGCATCTCTCTGCTCAGCTCCCAA 554
QY |||||
181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIlePheAsn 200
Db |||||
555 GTAGCTGGGACCAAGACATGCACCACTACACCTGCTGCTAATTTTATTTTATTTTAA 614
QY |||||
201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db |||||
615 TTTTGTGAGACAGAGTCTCAACTCTGTCAACCCAGGCTGGAGTGCAGTGGCGCAATCTTGGC 674
QY |||||
221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
Db |||||
675 TCACGTCAACCTCTGCTCCCGGTTCAAGTTATTTCTCTGCCAGGCTCTCTGAGTAGC 734
QY |||||
241 TrpAspTyrArgArgProArgLeuAlaAsnPhePheValPheLeuValGluMetGly 260
Db |||||
735 TGGGACTACAGGCGCCACACAGCCTAGCTAATTTTGTATTTTGTATTTTGTAGTAGAGTGGG 794
QY |||||
261 PheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeuProAlaSerAla 280
Db |||||
795 TTCACCATGTTGCCAGAGTGTATCTTGATCTCTGGACCTTGTGATCTGCTGCTGCTGCGGC 854
QY |||||
281 SerGlnSerAlaGlyIleThrGlyValSerHisHisAlaArgLeuIlePheAsnPheCys 300
Db |||||
855 TCCCAAGTGTGGGATTACAGCGCTGAGCCACCGCCCGCTTATTTTATTTTATTTTGT 914
QY |||||
301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320
Db |||||
915 TTGTTTGAATGGAATCTCACTCTGTATCCAGGCTGGAGTGAATGGCCAAATCTCGGC 974
QY |||||
321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340
Db |||||
975 TCACGTCAACCTCTGCTCCCGGCTCAAGCGATTTCTCTGTCTAGCTTCCCAAGCAGC 1034
QY |||||
341 TrpAspTyrGlyHisLeuProHisProAlaAsnPheCysIlePheIleArgGlyGly 360
Db |||||
1035 TGGGATTACGGGACCTGCCACCAACCCCGCTAATTTTGTATTTTGTATTTTGTATTTTGTATTTAGAGCGGG 1094
QY |||||
361 ValSerProTyrLeuSerGlyTrpSerGlnTrpProAspLeuArg 375
Db |||||
1095 GTTTCACCATATTTGTGAGGCTGGTCTCAACTCTGACCTCAGG 1139

```





```
|||||
15  ATGGAGTTTTCGCTCTTGTTGCCAGGCTGGAGTGCAATGGCGCAATCTCAGCTCACCGC 74
QY
21  AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
Db
75  AACCTCCGGCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCCCGAGTAGCTGGATT 134
QY
41  ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60
Db
135 ACAGGCATGTGCACCCACGCTCGGCTAATTTTGTATTTTATTTTATAGAGATGGAGTTT 194
QY
61  LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAla 80
Db
195 CTCATGTTGGTCAGGCTGGTCTCGAACTCCCGACTCAGATGATCCCTCGCTCGGCC 254
QY
81  SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys 100
Db
255 TCCCAAAGTGTAGATACAGGACTGGCCACCATGCGCGCTCTGCTGGCTAATTTTGT 314
QY
101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120
Db
315 CGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTCTGAGCTCAAGCAGTCCACC 374
QY
121 CysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheIle 140
Db
375 TGCCTCAGGCTCCCAAAGTGTGGGATACAGCGGTGCAGCGTGCCTGGCGCTTTTATT 434
QY
141 LeuPheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
Db
435 TTATTTTTTTTAAACACACAGGTGTCCCACTCTTACCCAGGATGAAGTGCAGTGGTGTAT 494
QY
161 HisSerSerLeuGlnProSerThrProGluIleLysHisProProAlaSerAlaSerGln 180
Db
495 CACAGCTCACTCAGCGTTTCAACTCCTTGAGATCAAGCATCTCTCTGCTCAGCCCTCCAA 554
QY
181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIlePheAsn 200
Db
555 GTAGCTGGGACCAAGACATGACACCTACACTGCTGGCTAATTTTATTTTATTTTAAAT 614
QY
201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db
615 TTTTGTGAGACAGAGTCTCAACTCTGTACCCAGGCTGGAGTGCAGTGGCGCAATCTTGGC 674
QY
221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
Db
675 TCACATGCAACCTCTCGCTCCCGGTTCAAGTTATTCTCTGCGCCAGCGCTCTCTGAGTAGC 734
QY
241 TrpAspTyrArgArgProProArgLeuAlaAsnPheValPheLeuValGluMetGly 260
Db
735 TGGGACTACAGGGGCCACACCGCTAGCTAATTTTGTATTTTATTTAGATAGATGGGG 794
QY
261 PheThrMetPheAlaArgLeuIleLeuSerGlyProCysAspLeuProAlaSerAla 280
Db
795 TTCAACATGTTCCGACAGGTGATCTTGATCTCTGGACCTTGATCTGCTGCTCGCTCGGCC 854
QY
281 SerGlnSerAlaGlyIleThrGlyValSerHisAlaArgLeuIlePheAsnPheCys 300
Db
855 TCCCAAAGTGTGGGATTACAGCGGTGAGCCACCGCCGCTTATTTTAAATTTTGT 914
QY
301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320
Db
915 TTGTTTGAATGGAATCTCACTCTGTACCCAGGCTGGAGTGCAATGGCCAAATCTCGGC 974
QY
321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340
Db
975 TCACATGCAACCTCTGCTCCCGGCTCAAGCGATTCTCTGCTCAGCTCCCAAGCAGC 1034
QY
341 TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGly 360
Db
1035 TGGGATTACGGGACCTGCACACACCCCGCTAATTTTGTATTTTATTTAGAGCGGG 1094
QY
361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375
|||||
```

```
Db
1095 GTTTCACCATATTGTCAGGCTGGTCTCAAACCTCTGACCTCAGG 1139
RESULT 7
AX709356 1442 bp DNA linear PAT 04-APR-2003
LOCUS
DEFINITION Sequence 1 from Patent WO02092115.
ACCESSION AX709356
VERSION AX709356.1 GI:29564893
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Averbach, P.A.
TITLE Method of preventing cell death using segments of neural thread
JOURNAL proteins
Patent: WO 02092115-A 1 21-NOV-2002;
NYMOX CORPORATION (CA)
FEATURES
source
Location/Qualifiers
1..1442
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
15..1142
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD88180.1"
/db_xref="GI:29564894"
/translation="MEFSLLLPLENCAGISAHNRLLPGSSDSPASASPVAGITGMC
THARLLLYFFLVEHFLHVGAGLELPTSDPSVSASQASARYRTHHARLCLANPCGR
NRSLMCPSPWSELPKQSLPKQMDYRAAVPGLFLLFRRHRCPTLTQDEOVNCD
HSSIQPSTPEIKHPKSPASQAVAGTKDMHYTWLIFLIFNFRSLNSVTOAGVQWRN
IGSLQPLPFPKLFSCPSLLSSWDYRRPPLANFFVLVFNFTFMFARLITISGFDL
PASASQASGITGVSHARLIENFCLFEMESHVTVQAGVQWPNLGLSLQPLPPLKRFSC
LSLPSWDYGHLPHPANFCIFIRGGVSPYLSGWSQTPDLR"
ORIGIN
Alignment Scores:
Pred. No.: 4,69e-142 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-380-203-2 (1-375) x AX709356 (1-1442)
QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
Db 15 ATGGAGTTTTCGCTCTTGTTGCCAGGCTGGAGTGCAATGGCGCAATCTCAGCTCACCGC 74
QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
Db 75 AACCTCCGGCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCCCGAGTAGCTGGATT 134
QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60
Db 135 ACAGGCATGTGCACCCACGCTCGGCTAATTTTGTATTTTATTTTATAGAGATGGAGTTT 194
QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAla 80
Db 195 CTCATGTTGGTCAGGCTGGTCTCGAACTCCCGACTCAGATGATCCCTCGCTCGGCC 254
QY 81 SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys 100
Db 255 TCCCAAAGTGTAGATACAGGACTGGCCACCATGCGCGCTCTGCTGGCTAATTTTGT 314
QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120
Db 315 CGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTCTGAGCTCAAGCAGTCCACC 374
QY 121 CysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheIle 140
```

```
Db 375 TGCCCTCAGCCTCCCAAGTGTGGGATTACAGCGTGCAGCGCTGCGCTTTTAAAT 434
QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
Db 435 TTATTTTTTTAAGACACACAGGTGTCACCTCTTACCAGGATGAAGTGCAGTGTGTGAT 494
QY 161 HisSerSerLeuGlnProSerThrProGluLeuLeuHisProProAlaSerAlaSerGln 180
Db 495 CACAGCTCACTGCAGCCTTCAACTCTGAGATCAAGCATCTCTCTGCGCTCAGCCTCCCAA 554
QY 181 ValAlaGlyThrLysAspMetHisHisTyThrTrpLeuLeuPheLeuPheLeuPheAsn 200
Db 555 GTAGCTGGGACCAAGACATGCACCACTACCTGGCTAATTTTTTATTTTTTAAAT 614
QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db 615 TTTTGTGAGACAGAGTCTCAACTCTGTACCCAGCTGGAGTGCAGTGGCGCAATCTTGCG 674
QY 221 SerLeuGlnProLeuProGlyPheLeuPheSerCysProSerLeuLeuSerSer 240
Db 675 TCACTCAACCTCTGCTCCCGGTTCAAGTTATTCTCTGCCGCCAGCCTCTGTAGTAGC 734
QY 241 TrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeuValGluMetGly 260
Db 735 TGGGACTACAGGCGCCACACAGCCTAGCTAATTTTTTGTATTTTGTAGTAGAGTGGGG 794
QY 261 PheThrMetPheAlaArgLeuLeuLeuSerGlyProCysAspLeuProAlaSerAla 280
Db 795 TTCACCATGTTCCCGAGGTGATCTTGATCTCGACCTTGTGATCTGCGCTCGCTCGGCC 854
QY 281 SerGlnSerAlaGlyIleThrGlyValSerHisHisAlaArgLeuLeuPheAsnPheCys 300
Db 855 TCCCAAAGTGTGGGATTACAGCGCGGAGCCACCCAGCGCGCTTATTTTAAATTTTGT 914
QY 301 LeuPheGluMetGlnSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320
Db 915 TTGTTTGAATGGAAATCTCACTCTGTATCCAGCGCTGGAGTGAATGGGCAAAATCTCGGC 974
QY 321 SerLeuGlnProLeuProGlyLeuLysAspGlyPheSerCysLeuSerLeuProSerSer 340
Db 975 TCACCTCAACCTCTGCTCCCGGCTCAAGCGATTCTCTGTCTCAGCCTCCCAAGCAGC 1034
QY 341 TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGlyGly 360
Db 1035 TGGGATTACGGGCACCTGCCACACACCCCGCTAATTTTGTATTTTTCATTAGAGCGGG 1094
QY 361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375
Db 1095 GTTTCACCATATTTGTACAGCTGGTCTCAAACTCTCAGCCTCAGG 1139
```

## RESULT 8

```
AX770807 1442 bp DNA linear PAT 02-JUL-2003
LOCUS AX770807
DEFINITION Sequence 125 from Patent WO03008444.
ACCESSION AX770807
VERSION AX770807.1 GI:32437975
```

## SOURCE

```
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

## REFERENCE

```
1
Averback, P.A. and Gemmell, J.
Peptides effective in the treatment of tumors and other conditions
requiring the removal or destruction of cells
Patent: WO 03008444-A 125 30-JAN-2003;
NYMOX CORPORATION (CA)
```

## FEATURES

```
Location/Qualifiers
1..1442
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

## CDS

```
15..1142
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE01151.1"
/db_xref="GI:32437976"
/translation="MEFSLLLPRLECNCAISAHNRLLRPGSSDSDSPASAPFVAGITGMC
THARLLLYFFLEMEFLHVQAGLELPTSDPSVASOSARVYRTHGHARLCIANFCGR
NRVSLMCPSPPELKQSTCLSLPKCWDRRAAVPGFLFLFRHRCPTLTQDEQWRN
HSSLOPSTPEIKHPASASQVAGTDMHHTWLIFFIFNFRQSLNSVTQAGVQRN
LGSLOPLPGPKFLFSCPSLLSSWDYRRPRLANFFVFLFMGTFTMPARILLISGPDCL
PASASOSAGITGVSHHARLIENFCLFEMESHVSVTQAGVQWPNLGSLOPLPGLKRFSC
LSLPSSWDYGHLPHPANFCIFIRGVSFYLGSWSQTPDLR"
```

## ORIGIN

## Alignment Scores:

```
Pred. No.: 4 69e-142 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
```

US-09-380-203-2 (1-375) x AX770807 (1-1442)

```
QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
Db 15 ATGGAGTTTTTGGCTCTTGTGTCAGGCTGGAGTGCATATGGCGCAATCTCAGCTCACC 74
QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
Db 75 AACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCTCAGCCTCCCGCTAGTGGGATTT 134
QY 41 ThrGlyMetCysThrHisAlaArgLeuLeuLeuTyPhePheLeuValGluMetGluPhe 60
Db 135 ACAGCATGTGACCCACCGCTCGGCTAATTTTGTATTTTTTTTAGTAGAGTGGAGTTT 194
QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAla 80
Db 195 CTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATCCCTCCGCTCGGCC 254
QY 81 SerGlnSerAlaArgTyrArgThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCys 100
Db 255 TCCCAAGTGTAGATACAGGACTGGCCACCATGCGCGCTCTGCGCTGGCTAATTTTGT 314
QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120
Db 315 GGTAGAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTCTGAGCTCAAGCAGTCCACC 374
QY 121 CysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheIle 140
Db 375 TGCCTCAGCCTCCCAAGTGTGGGATTACAGCGGTGCAGCGTGCCTGGCTTTTAAAT 434
QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
Db 435 TTATTTTTTTAAGACACACAGGTGTCACCTCTTACCAGGATGAAGTGCAGTGGTGTGAT 494
QY 161 HisSerSerLeuGlnProSerThrProGluLeuLysHisProProAlaSerAlaSerGln 180
Db 495 CACAGCTCACTGCAGCCTTCAACTCTGTAGATCAAGCATCTCTCTGCGCTCAGCCTCCCAA 554
QY 181 ValAlaGlyThrLysAspMetHisHisTyThrTrpLeuLeuPheLeuPheLeuPheAsn 200
Db 555 GTAGCTGGGACCAAGACATGCACCACTACACCTGGCTAATTTTTTATTTTTTAAAT 614
QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db 615 TTTTGTGAGACAGAGTCTCAACTCTGTACCCAGCTGGAGTGCAGTGGCGCAATCTTGCG 674
QY 221 SerLeuGlnProLeuProGlyPheLeuPheSerCysProSerLeuLeuSerSer 240
Db 675 TCACTGCAACCTCTGCTCCCGGTTCAAGTTATTCTCTGCCGCCAGCCTCTGTAGTAGC 734
QY 241 TrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeuValGluMetGly 260
```

```

Db      735  TGGGACTACAGGCGCCACACCGCCTAGCTAATTTTTTTGTATTTTAGTAGAGATGGGG 794
Qy      261  PheThrMetPheAlaArgLeuLeuLeuLeuSerGlyProCysAspLeuProAlaSerAla 280
Db      795  TTCAACCATGTCGCCAGGTGATCTTGTGATCTCTGGACCTTGTGATCTGCTCGCTCGGCC 854
Qy      281  SerGlnSerAlaGlyIleThrGlyValSerHisHisAlaArgLeuLeuLeuPheAsnPheCys 300
Db      855  TCCCAAGTGTGGGATACAGCGGTGAGCCACCGCCCGGCTATTTTTTAATTTTGT 914
Qy      301  LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320
Db      915  TTGTTTGAATGGAATCTCACTCTGTACCCAGGCTGGAGTGCATGGCCAAATCTCGGC 974
Qy      321  SerLeuGlnProLeuProGlyLeuLeuLeuArgPheSerCysLeuSerLeuProSerSer 340
Db      975  TCACTGCAACCTCTGCTCCCGGCTCAAGCGATTCTCTCTCAGCCTCCCAAGCAGC 1034
Qy      341  TrpAspTyrGlyHisLeuProHisProAlaAsnPheCysIlePheIleArgGlyGly 360
Db      1035  TGGGATTAGCGGCACCTGCCACACACCGCCGCTAATTTTGTATTTTCAATTAGAGCGGG 1094
Qy      361  ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375
Db      1095  GTTTCACCATATTTGTTCAGCGTGGTCTCAAACTCTCAGCCTCAGG 1139

RESULT 9
AX774635
LOCUS      1442 bp      DNA      linear      PAT 09-JUL-2003
DEFINITION      Sequence 53 from Patent WO02097030.
ACCESSION      AX774635
VERSION      AX774635.1  GI:32486166
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS      Averbach, P.A.
TITLE      Peptides derived from neural thread proteins and their medical use
JOURNAL      Patent: WO 02097030-A 53 05-DEC-2002;
NYMOX CORPORATION (CA)
FEATURES
source
1..1442
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
15..1142
/codon_start=1
/product="AD7C-NTP"
/protein_id="CAE06497.1"
/db_xref="GI:32486167"
/translation="MEFSLLPRLCNGAISAHNRRLRPGSSDSPASAPVAGITGMC
THARILYFVLEMFHVGAGLELPTSDPSVASOSARYRTGHARLCLANFCGR
NRVSLMCPWSPELQSKSLCLPKCDWDRRAVPGFLFLFLRHRCPILTQDEVQWCD
HSSLQPSPEIKHPASASQVAGTKDMHYTWLIFIFLQRSLNSVTQGVQWRN
LGSLOPLPGFLKSCPLLSLSDWRPRPLANFVFLVEMGFTWPARLILISGPCDL
PASQSQSAGITGVSHARLIFNFCFLFEMESHSVTQGVQWPNLGLSLQPLPPLKRFSC
LSLPSWSDYGHLPHPHPANFCIFIRGVSPYLSGWSQTPDLR"
ORIGIN
Alignment Scores:
Pred. No.:      4.69e-142      Length:      1442
Score:      2034.00      Matches:      375
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      6      Gaps:      0
US-09-380-203-2 (1-375) x AX774635 (1-1442)
Qy      1  MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20

```

```

Db      15  ATGGAGTTTTCGCTCTTGTGTGCCAGGCTGGAGTGCATATGGCGCAATCTCAGCTCACCGC 74
Qy      21  AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
Db      75  AACCTCCGCTCCCGGGTTCAAGCGATTCTCTCGCTCAGCCTCCAGTAGTGGATT 134
Qy      41  ThrGlyMetCysThrHisAlaArgLeuLeuLeuTyrPhePheLeuValGluMetGluPhe 60
Db      135  ACAGCATGTGCACCGCTCGGCTAATTTTGTATTTTATTTTATTTAGTAGAGATGGAGTT 194
Qy      61  LeuHisValGlnAlaGlyLeuLeuProThrSerAspSerProSerValSerAla 80
Db      195  CTCCATGTGTGTGAGGCTGGTCTCGAATCCCGACCTCAGATGATCCCTCGCTCTCGGCC 254
Qy      81  SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys 100
Db      255  TCCCAAGTGTAGATACAGGACTGGCCACCATGCCCGGCTCTGCTGGCTAATTTTGT 314
Qy      101  GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120
Db      315  GGTAGAAAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTCTGAGCTCAAGCAGTCCACC 374
Qy      121  CysLeuSerLeuProLysCysTrpAspTyrArgAlaAlaValProGlyLeuPheIle 140
Db      375  TGCTCTCAGCCTCCCAAGTGTGGGATTACAGGCGTGCAGCGTGCCTGGGCTTTTATT 434
Qy      141  LeuPheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
Db      435  TTATTTTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAGTCAGTGGTGTGAT 494
Qy      161  HisSerSerLeuGlnProSerThrProGluLeuLysHisProProAlaSerAlaSerGln 180
Db      495  CACAGCTCACTGCAGCCTTCAACTCTCGAGATCAAGCATCTCTCTGCTGCTCAGCCTCCCAA 554
Qy      181  ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIlePheIlePheAsn 200
Db      555  GTAGCTGGGACCAAGACATGCACCATCACCTGGCTAATTTTTTATTTTATTTTAAAT 614
Qy      201  PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db      615  TTTTGTGACACAGAGTCTCAACTCTGTCAACCAGGCTGGAGTGCAGTGGCGCAATCTTGGC 674
Qy      221  SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
Db      675  TCACGTGCNAACCTCTGCTCCCGGGTTCAAGTATTTCTCTCTCCCGCAGCCTCTCTGAGTAGC 734
Qy      241  TrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeuValGluMetGly 260
Db      735  TGGGACTACAGGCGCCACACCGCCTAGCTAATTTTTTTGTATTTTATTTAGTAGAGATGGGG 794
Qy      261  PheThrMetPheAlaArgLeuLeuLeuSerGlyProCysAspLeuProAlaSerAla 280
Db      795  TTCACATGTTTCGCGAGGTGATCTTGTATCTCTGGACCTTGTGATCTGCTCGCTCGGCC 854
Qy      281  SerGlnSerAlaGlyIleThrGlyValSerHisHisAlaArgLeuLeuPheAsnPheCys 300
Db      855  TCCCAAGTGTGGGATTTACAGGCGTGCAGCCACCGCCCGGCTATTTTAAATTTTGT 914
Qy      301  LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320
Db      915  TTGTTTGAATGGAATCTCACTCTGTACCCAGGCTGGAGTGCATGGCCAAATCTCGGC 974
Qy      321  SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340
Db      975  TCACTGCAACCTCTGCTCCCGGCTCAAGCGATTCTCTCTCAGCCTCCCAAGCAGC 1034
Qy      341  TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGlyGly 360
Db      1035  TGGGATTAGCGGCACCTGCCACACACCGCCGCTAATTTTGTATTTTCAATTAGAGCGGG 1094
Qy      361  ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375

```

```

Db      1095 GTTTCACCATATTTGTGACGGCTGGTCTCAAACTCTCTGACCTCAGG 1139
RESULT 10
LOCUS   BD070526                      1442 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION Transgenic animals and cell lines for screening drugs effective for the treatment or prevention of Alzheimer's disease.
ACCESSION BD070526
VERSION   BD070526.1 GI:22616129
KEYWORDS  JP 2001513777-A/1.
SOURCE   unidentified
ORGANISM  unclassified.
REFERENCE 1 (bases 1 to 1442)
AUTHORS   Monte,S.D. and Wands,J.R.
TITLE     Transgenic animals and cell lines for screening drugs effective for the treatment or prevention of Alzheimer's disease
JOURNAL   Patent: JP 2001513777-A 1 04-SEP-2001;
          THE GENERAL HOSPITAL CORP
COMMENT   OS   Unidentified
          PN   JP 2001513777-A/1
          PD   04-SEP-2001
          PR   26-FEB-1998 JP 1998537813
          PR   26-FEB-1997 US 60/038908
          PI   SUZANNE DELA MONTE,JACK R WANDS
          PC   C07H21/02,C07H21/04,C12N5/00,C12N15/00,C12Q1/02,A61K48/00, PC A61K49/00
          CC   Strandedness: Double;
          CC   Topology: Both;
          CC   Transgenic animals and cell lines for screening drugs CC
          CC   effective for the treatment or prevention of Alzheimer's disease FH Key
          CC   Location/Qualifiers
          FT   CDS      15..1139.
              Location/Qualifiers
              1..1442
              /organism="unidentified"
              /mol_type="genomic DNA"
              /db_xref="taxon:32644"

ORIGIN
Alignment Scores:
Pred. No.:      4,69e-142      Length:      1442
Score:          2034.00      Matches:      375
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             6             Gaps:        0

US-09-380-203-2 (1-375) x BD070526 (1-1442)

Qy      1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyValAlaHisArg 20
Db      15 ATGGAGTTTTCGCTTGTTGCCAGCTGGAGTGCATGGCGCAATCTCAGCTACCCG 74

Qy      21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
Db      75 AACCTCCGGCTCCCGGGTTCAAGCGATTCTCTGCTCAGCTCCCGCTCAGCTGGGATT 134

Qy      41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60
Db      135 ACAGGCATGTGCACCCACCTCGGCTAAATTTTGTATTTTGTATTTTGTAGATGGAGTTT 194

Qy      61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAla 80
Db      195 CTCCTATGTTGGTCAGCTGGTCTCGAACTCCCGACCTCAGATGATCCCTCCGCTCGGCC 254

Qy      81 SerGlnSerAlaArgTyrArgThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCys 100
Db      255 TCCCAAAGTGTAGATACAGGACTGGCCACCATGCGCGCTCTGCTGCTGCTAAATTTTGT 314

Qy      101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerTrpSerProGluLeuLysGlnSerThr 120

```

```

Db      315 GGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGCTCAAGCAGTCCACC 374
Qy      121 CysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheIle 140
Db      375 TGCCCTCAGCTCCCAAAGTGTGGGATTACAGCGGTGACAGCGTGCCTGGCTTTTATT 434
Qy      141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
Db      435 TTATATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAGTGCAGTGGTGTGAT 494
Qy      161 HisSerSerLeuGlnProSerThrProGluLysHisProAlaSerAlaSerGln 180
Db      495 CACAGCTCACTGCAGCTTCAACTCTCTGAGATCAAGCATCTCTGCTCAGCTCCCA 554
Qy      181 ValAlaGlyThrLysAspMethHisIleTyrThrTrpLeuIlePheIlePheAsn 200
Db      555 GTAGCTGGGACCAAGACATGCACCATACCTGCTGCTAAATTTTATTTTATTTTAAAT 614
Qy      201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db      615 TTTTGTAGACACAGTCTCAACTCTGTCAACCCAGGCTGGAGTGCAGTGGCGCAATCTTGGC 674
Qy      221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
Db      675 TCACCTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCCAGCTCTCTGAGTAGC 734
Qy      241 TrpAspTyrArgArgProProArgLeuAlaAsnPheValPheLeuValGluMetGly 260
Db      735 TGGGACTACAGCGGCCACCGCTAGCTAAATTTTGTATTTTGTATTTTGTAGATGGGG 794
Qy      261 PheThrMetPheAlaArgLeuIleLeuSerGlyProCysAspLeuProAlaSerAla 280
Db      795 TTCACCATGTTGCCCAGGTTGATCTTGATCTCTGGACCTTGTGATCTGCTCGCTCGGCC 854
Qy      281 SerGlnSerAlaGlyIleThrGlyValSerHisHisAlaArgLeuIlePheAsnPheCys 300
Db      855 TCCCAAAGTGTGGGATTACAGCGCTGAGCCACCAACCGCGCTTATTTTAAATTTTGT 914
Qy      301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320
Db      915 TTGTTTGAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCATGGCAATCTCGGC 974
Qy      321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340
Db      975 TCACCTGCAACCTCTCGCTCCCGGGCTCAAGCGATTCTCTGCTCAGCTCCCAAGCAGC 1034
Qy      341 TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGlyGly 360
Db      1035 TGGGATTACGGGCACCTGCCACACACCCCGCTAAATTTTGTATTTTTCATTAGAGCGGG 1094
Qy      361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375
Db      1095 GTTTCACCATATTTGTGAGCTGGTCTCAAACTCTCTGACCTCAGG 1139

```

```

RESULT 11
LOCUS   AF010144                      1442 bp      mRNA      linear      PRI 02-NOV-2001
DEFINITION Homo sapiens neuronal thread protein AD7c-WTP mRNA, complete cds.
ACCESSION AF010144
VERSION   AF010144.1 GI:3002526
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 1442)
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE     Monte,S.M., Ghanbari,K., Frey,W.H., Beheshti,I., Averbach,P., Hauser,S.L., Ghanbari,H.A. and Wands,J.R.
          Characterization of the AD7c-WTP cDNA expression in Alzheimer's disease and measurement of a 41-kD protein in cerebrospinal fluid
JOURNAL   J. Clin. Invest. 100 (12), 3093-3104 (1997)
MEDLINE   98064067
PUBMED    9399956

```



```

REFERENCE
AUTHORS de la Monte,S.M. and Wands,J.R.
TITLE Direct Submission
JOURNAL Submitted (24-JUN-1997) MGH Cancer Center, Massachusetts General
Hospital, 149 13th Street, Room 7308, Charlestown, MA 02129, USA
FEATURES
Source
1..1442
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="neuronal"
15..1142
/notes="contains expressed Alu-containing domains; 41 kDa
phosphoprotein; cell death and neuritic sprouting observed
in neuronal cells transfected with the AD7c-NTP cDNA"
/codon_start=1
/product="neuronal thread protein AD7c-NTP"
/protein_id="AAC08737.1"
/db_xref="GI:3002527"
/translation="MEFSLLPRLCNGAISAHNRILRLPGSSDSPASAPVAGITGMC
THARLILYFLVEMEFHVGQAGLEBLPTSDPSVASOSARYTGHARLCLANFCGR
NRVSLMCPSPWELQSTCLSLPKCDYRRAAVPGLFIPLFHRHRCPTLTQDEVQWCD
HSSLOPSTPEIKHPSPASQVAGTKDMHHTWLIIFIFNFRQSLNSVTQAGVQWRN
LGSLOPLPGFKLFCPSLLSDWYRRPRLANFFVFLVMGFTWPARLILISGPCDL
PASASQAGITGVSHHARLIFNFCLEFEMESHVSTQAGVQWPNLGSLLQPLPPGLKRFSC
LSLPSSWDYGHLPHPANFCIFIRGVSPYLSGNSQTPDLR"
ORIGIN
Alignment Scores:
Pred. No.: 4,69e-142 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-380-203-2 (1-375) x AF010144 (1-1442)
QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
Db 15 ATGGAGTTTTTCGGCTCTGTGGCCAGGCTGGAGTGCATGGCGCAATCTCAGCTCAGCCG 74
QY 21 AnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
Db 75 AACCTCGGCTCCCGGGTTCAAGCAATCTCTCGCTCAGCTCCCGCTCAGTGTGGATT 134
QY 41 ThrGlyMetCysThrHisAlaArgLeuLeuLeuTyrPhePheLeuValGluMetGluPhe 60
Db 135 ACAGGCATGTGCACCCACGCTCGGCTAAATTTGTTATTTTGTAGTAGAGATGGAGTTT 194
QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAla 80
Db 195 CTCCTCATGTTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCCCTCGCTCGGCC 254
QY 81 SerGlnSerAlaArgTyrArgThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCys 100
Db 255 TCCCAAGTGCTAGATACAGGACTGGCCACCATGCGCGCTCTCGCTCGGCTAAATTTTGT 314
QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerTrpSerProGluLeuLysGlnSerThr 120
Db 315 GGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGCTCTCCTCAGCTCAAGCAGTCCACC 374
QY 121 CysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheIle 140
Db 375 TGCTCAGGCTCCCAAGTGTGGGATTCAGCGGTGCAGCGGTGCCTCGGCTTTTATT 434
QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
Db 435 TTATTTTTTTTAAAGACACAGGTGTCCTACTCTTACCAGGATGAAGTGCAGTGGTGAT 494
QY 161 HisSerSerLeuGlnProSerThrProGluIleLysHisProProAlaSerAlaSerGln 180
Db 495 CACAGCTCACTGAGGCTTTCAACTCTGAGATCAAGCATCTCTCGCTCAGCTCCCAA 554

```

```

QY 181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIlePheAen 200
Db 555 GTAGCTGGGACCAAGACATGCACACATACACCTGGCTAAATTTTATTTTAAAT 614
QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db 615 TTTTGTAGACAGAGTCTCAACTCTGTCCACGAGCTGGAGTGCAGTGGCGCAATCTTGGC 674
QY 221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
Db 675 TCACGTGAACCTCTCGCTCCCGGTTCAAGTTATTTCTCTGCTCCCGCAGCTCTGAGTAGC 734
QY 241 TrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeuValGluMetGly 260
Db 735 TGGGACTACAGCGGCCACCGCTAGCTAATTTTGTATTTTGTATTTAGTAGAGATGGGG 794
QY 261 PheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeuProAlaSerAla 280
Db 795 TTCACATGTTCCGCAAGTTGATCTTGATCTCTGACCTTGATCTCGCTCGCTCGGCC 854
QY 281 SerGlnSerAlaGlyIleThrGlyValSerHisHisAlaArgLeuIlePheAsnPheCys 300
Db 855 TCCCAAGTGTCTGGATTACAGGCTGAGCCACCGCCGGCTTATTTTAAATTTGT 914
QY 301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320
Db 915 TTGTTTGAATGGAATCTCACTCTGTATCCCGAGCTGGAGTGCATGGCCAAATCTCGGC 974
QY 321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340
Db 975 TCACGTGAACCTCTCGCTCCCGGCTCAAGCAATCTCTGCTCAGCTCCCAAGCAGC 1034
QY 341 TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGlyGly 360
Db 1035 TGGGATTACGGGCACCTGCCCACACCGCCGTAATTTTGTATTTTCAATTAGAGCGGG 1094
QY 361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375
Db 1095 GTTTCACCATATTTGTGAGGCTGGTCTCAAACTCTCGACCTCAGG 1139
RESULT 12
HS886K2/c
LOCUS
DEFINITION
Human DNA sequence from clone RP5-886K2 on chromosome 1p35.1-36.12
Contains elongin A (RNA polymerase II elongation factor),
hydroxymethylglutaryl-CoA lyase, GALE (UDP-galactose-4-epimerase)
ESTs, GSS, STS, CpG island, complete sequence.
ACCESSION
AL031295
VERSION
AL031295.1 GI:4376011
KEYWORDS
HTG; elongin A; GALE; HMGCL; hydroxymethylglutaryl-CoA lyase; RNA
polymerase II elongation factor.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 124001)
Frankland,J.
Direct Submission
Submitted (05-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Mar 7, 1999 this sequence version replaced gi:4056542.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission

```



corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human Chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr1>

RP5-886K2 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pCIPAC2

This sequence is the entire insert of clone RP5-886K2.

#### FEATURES

##### source

Location/Qualifiers  
1..124001

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="RZPD:RPCIP704K02886"

/db\_xref="taxon:9606"

/chromosomes="1"

/map="p35.1-36.12"

/clone="RP5-886K2"

/clone\_lib="RPCI-5"

50..160

/note="AluY repeat: matches 195..302 of consensus"

161..324

/note="L1M5 repeat: matches 6133..6297 of consensus"

327..594

/note="AluJb repeat: matches 1..294 of consensus"

597..863

/note="AluJo repeat: matches 5..298 of consensus"

868..1170

/note="AluYb8 repeat: matches 1..310 of consensus"

1671..1848

/note="MIR repeat: matches 67..261 of consensus"

2288..2429

/note="AluJo/FRAM repeat: matches 157..300 of consensus"

2430..2714

/note="AluSq repeat: matches 1..287 of consensus"

2728..3063

/note="L1M4 repeat: matches 2200..2535 of consensus"

3385..3686

/note="AluSx repeat: matches 1..310 of consensus"

3709..4050

/note="THE1C repeat: matches 1..371 of consensus"

4064..4212

/note="FRAM repeat: matches 1..155 of consensus"

4358..4506

/note="MER5A repeat: matches 15..182 of consensus"

4788..4878

/note="AluJo repeat: matches 54..149 of consensus"

4879..5178

/note="AluSx repeat: matches 1..302 of consensus"

5181..5484

/note="AluSg repeat: matches 1..305 of consensus"

5485..5653

/note="AluJo repeat: matches 149..304 of consensus"

5654..5788

/note="AluSx repeat: matches 1..132 of consensus"

5789..6089

/note="AluSp repeat: matches 1..297 of consensus"

6090..6260

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

/note="AluSx repeat: matches 132..312 of consensus"

6282..6580

/note="AluY repeat: matches 1..296 of consensus"

6596..7224

/note="L2 repeat: matches 1995..2748 of consensus"

7225..7293

/note="AluJb repeat: matches 152..220 of consensus"

7294..7753

/note="L2 repeat: matches 1543..1995 of consensus"

7756..8048

/note="AluSg repeat: matches 1..279 of consensus"

8050..8358

/note="AluY repeat: matches 1..309 of consensus"

8420..8987

/note="L2 repeat: matches 873..1539 of consensus"

8998..9090

/note="L2 repeat: matches 1543..1627 of consensus"

9060..9105

/note="L2 repeat: matches 827..872 of consensus"

9180..9279

/note="MIR repeat: matches 35..139 of consensus"

9489..9671

/note="MER5A repeat: matches 1..189 of consensus"

10530..10594

/note="L2 repeat: matches 2676..2750 of consensus"

10604..10678

/note="MIR repeat: matches 12..87 of consensus"

11227..11408

/note="MIR repeat: matches 20..232 of consensus"

11470..11776

/note="AluSg repeat: matches 1..311 of consensus"

12244..12546

/note="AluY repeat: matches 1..303 of consensus"

13021..13318

/note="AluSx repeat: matches 1..295 of consensus"

13319..13389

/note="MER5A repeat: matches 73..137 of consensus"

13390..13685

/note="AluSg repeat: matches 1..297 of consensus"

13686..13730

/note="MER5A repeat: matches 27..73 of consensus"

13759..13823

/note="L2 repeat: matches 2146..2215 of consensus"

13824..13959

/note="FLAM A repeat: matches 1..133 of consensus"

13960..14115

/note="L2 repeat: matches 2012..2146 of consensus"

14221..14479

/note="MER21B repeat: matches 156..456 of consensus"

14660..14709

/note="MIR repeat: matches 49..98 of consensus"

14953..15263

/note="AluSx repeat: matches 1..311 of consensus"

15508..15684

/note="MIR repeat: matches 2..218 of consensus"

15724..15857

/note="FLAM A repeat: matches 1..133 of consensus"

15961..16086

/note="FAM repeat: matches 12..163 of consensus"

16087..16394

/note="AluSg repeat: matches 3..310 of consensus"

16395..16409

/note="FAM repeat: matches 163..174 of consensus"

16722..16804

/note="L2 repeat: matches 2582..2691 of consensus"

16908..17483

/note="L1MB3 repeat: matches 5598..6185 of consensus"

17679..17881

/note="L2 repeat: matches 2285..2490 of consensus"

18516..18927

/note="MLTID repeat: matches 114..505 of consensus"

18928..19043

/note="FLAM\_C repeat: matches 1..116 of consensus"

```

repeat_region 19044..19149
/note="MUTID repeat: matches 6. .114 of consensus"
repeat_region 19750..20088
/note="L2 repeat: matches 2273. .2646 of consensus"
repeat_region 20092..20153
/note="31 copies 2 mer aa 80% conserved"
repeat_region 20170..20479
/note="AluX repeat: matches 3. .312 of consensus"
repeat_region 20499..20633
/note="L2 repeat: matches 2623. .2750 of consensus"
repeat_region 20639..20946
/note="AluX repeat: matches 1. .308 of consensus"
repeat_region 20959..21232
/note="AluX repeat: matches 1. .297 of consensus"
repeat_region 21240..21532
/note="AluSg repeat: matches 1. .293 of consensus"
repeat_region 21534..21837
/note="AluSp repeat: matches 2. .310 of consensus"
repeat_region 21839..21936
/note="WERSB repeat: matches 84. .175 of consensus"
repeat_region 21952..22283
/note="L2 repeat: matches 1811. .2198 of consensus"
repeat_region 22303..22605
/note="AluY repeat: matches 1. .301 of consensus"

Alignment Scores:
Pred. No.: 2,63e-122 Length: 124001
Score: 1799.00 Matches: 360
Percent Similarity: 94.52% Conservative: 2
Best Local Similarity: 93.99% Mismatches: 12
Query Match: 88.45% Indels: 11
DB: 9 Gaps: 2

US-09-380-203-2 (1-375) x HS886K2 (1-124001)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyValAlaIleSerAlaHisArg 20
DB 6072 ATGGAGTTTCGCTCTGTGTCCAGGCTGGAGTGCATGGCGCAATCTCAGCTCACGCC 6013

QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
DB 6012 AACCTCCGCTCCCGGGTTCAGCGATTCTCTGCTCAGCTCCCTAGTAGCTGGATT 5953

QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60
DB 5952 ACAGGATGTGCCACACCGCTCGGCTAAATTTGTATTTTGTATTTTGTAGATGGAGTT 5893

QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspSerProSerValSerAla 80
DB 5892 CTCATGTTGGTCAGGCTGGTCTCGAATCCCGACCTCAGATGATCCTCCGCTCGGCC 5833

QY 81 SerGlnSerAlaArgTyrArgThrGly-----HisHisAlaArg-LeuCysLeuAlaAs 98
DB 5832 TCCCAAAGTGTGAGATT---ACAGGCATGAGCCACCATGCGCGGCTCTGCTGGCTAA 5776

QY 98 nPheCysGlyArgAsnArgValSerLeuMet-CysProSerTrpSerProGluLeuLysG 118
DB 5775 TTTTGTGTGTAAGAACAGGGTTTCACTGATGTGGCCAAAGCTGGTCTCTGAGCTCAAGC 5716

QY 118 LnsSerThrCysLeuSerLeuProLysCysTrpAspTyrArgArg--AlaAlaValProGl 137
DB 5715 AGTCCACCTGCCTCAGCTCCCAAGTGTGGGATTCAGGCGTGAGCCACCGTGCCTGG 5656

QY 137 yLeuPheLeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGl 157
DB 5655 CTTTTTATTTATTTTATTTTAAAGACACAGGTGTCCACTCTTACCGAGATGAAGTGCA 5596

QY 157 nTrpCysAspHisSerSerLeuGlnProSerThrProGluIleLys-HisProProAlaAs 177
DB 5595 GTGGTGTATCAGCTACTCAGGCTTCACTCTGAGATCAAGCAATCTCTCGCT 5536

QY 177 erAlaSerGlnValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheLeuP 197

```

```

Db 5535 CAGCTCCCAAGTAGCTGGGACCAAGACATGCACCACCTACACCTGGCTAATTTTTATT 5476
QY 197 heIlePheAsnPheLeuArgGlnSerLeuAanSerValThrGlnAlaGlyValGlnTrpA 217
Db 5475 TTAATTTAAATTTTGTAGACAGAGTCTCAC-TCTGTCCACCGAGCTGGAGTGCAGTGC 5417
QY 217 rgAsnLeuGlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerL 237
Db 5416 GCAATTTGGGTCACTGCAACCTCTGCTCCGGGTTCAGGTATTTCTCTGCCCCAGCC 5357
QY 237 euLeuSerSerTrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeuV 257
Db 5356 TCCTGAGTAGCTGGGACTACAGGCGCCACACCGCTAGCTAATTTTTTGTATTTTAG 5297
QY 257 alGluMet-GlyPheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeu 276
Db 5296 TAGAGATGGGGTTTCACCATGTTCCAGGTGGTCTTGATCTCTGACCTTGTGATCTG 5237
QY 277 ProAlaSerAlaSerGlnSerAlaGlyIleThrGlyValSerHisHisAlaArgLeuIle 296
Db 5236 CTGCTCGGCTCCCAAAGTGTGGGATTACAGCGTGAGCCACCGCGGCTATT 5177
QY 297 PheAsnPheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrp 316
Db 5176 TTTAATTTTGTGTTGAAATGGAATCTCCTCTGTACCGAGCTGGAGTGCATGG 5117
QY 317 ProAsnLeuGlySerLeuGlnProLeuProProGlyLeuLysArgPheSerCysLeuSer 336
Db 5116 CACAATCTCGGCTCACTGCAACCTCTGCTCCCGGTTCAAGCGATTCTCTGTCTCAGC 5057
QY 337 LeuProSerSerTrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePhe 356
Db 5056 CTCCCAAGCAGCTGGGATTACGGGCACCTGCCAA--CACCCCGCTAAATTTTGTATTTTC 4999
QY 357 IleArgGlyGlyValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375
Db 4998 ATTAGAGCGGGGTTTCAACATATTTGTGAGGCTGTCTCAAACTCTCAACCTCCTCAG 4942

RESULT 13
BD070528 1418 bp DNA linear PAT 27-AUG-2002
LOCUS Transgenic animals and cell lines for screening drugs effective for
DEFINITION the treatment or prevention of Alzheimer's disease.
ACCESSION BD070528
VERSION BD070528.1 GI:22616131
KEYWORDS JP 2001513777-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1418)
AUTHORS Monte,S.D. and Wands,J.R.
TITLE Transgenic animals and cell lines for screening drugs effective for
JOURNAL the treatment or prevention of Alzheimer's disease
PATENT: JP 2001513777-A 3 04-SEP-2001;
THE GENERAL HOSPITAL CORP
COMMENT OS Unidentified
PN JP 2001513777-A/3
PD 04-SEP-2001
PF 26-FEB-1998 JP 1998537813
PR 26-FEB-1997 US 60/038908
PI SUZANNE DELA MONTE, JACK R WANDS
PC C07H21/02, C07H21/04, C12N5/00, C12N15/00, C12Q1/02, A61K48/00, PC
A61K49/00
CC Strandedness: Both;
CC Topology: Both;
CC Transgenic animals and cell lines for screening drugs CC
CC effective for the
CC treatment or prevention of Alzheimer's disease FH Key
FT Location/Qualifiers
FT source 1. .1418
/organism='Unidentified'.
Location/Qualifiers
1. .1418
source

```

/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.08e-123 Length: 1418  
Score: 1787.00 Matches: 363  
Percent Similarity: 95.54% Conservativeness: 11  
Best Local Similarity: 95.28% Mismatches: 1  
Query Match: 87.86% Indels: 10  
DB: 6 Gaps: 0

US-09-380-203-2 (1-375) x BD070528 (1-1418)

```
QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
DB 14 ATGGAGTTTTCGCTCTTTGTCAGGCTGGAGTGCATGGCGCAATCTCAGCTCACCGC 73
QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProVal-AlaGlyI 40
DB 74 AACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCGCTCAGCCTCCCGAGTAGCTGGAT 133
QY 40 eThrGlyMetCysThrHisAlaArgLeuLeuTyrPhePheLeuValGluMetGluPh 60
DB 134 TACAGGCATGTGCAC-CACGCTCGGCTAAATTTTGTATTTTATTTAGTAGAGTGAGTT 192
QY 60 eLeuHisValGlyGlnAlaGlyLeuLeuProThrSerAspAspProSerValSerAl 80
DB 193 TCTCCATGTTGTCAGGCTGGTCTCGAATCTCCGACCTCAGATGATCTCCCGCTTCGGC 251
QY 80 aSerGlnSerAlaArgTyrArgThrGlyHisAlaArg-LeuCysLeuAlaAsnPhC 100
DB 252 CTCCTCCAAAGTGTAGATACAGACTGAGCACCATGTCGCGGCTCTGCTGGCTAATTTT 311
QY 100 yGlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerT 120
DB 312 GTGGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTCGAGCTCAAGCAGTCCA 371
QY 120 hCysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheI 140
DB 372 CTGCTCAGCTCCCGGCTCCCGGCTGAGGATGACAGGCTGAGGCTGAGGCTGAGGCTTTTA 431
QY 140 leLeuPheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysA 160
DB 432 TTTTATTTTATTAAGACACAGGTGTCCCACTCTTACCAGGATGAAGTGCAGTGGTGTG 491
QY 160 sPheSerSerLeuGlnProSerThrProGluLeuLysHisProAlaSerAlaSerG 180
DB 492 ATCAGAGCTCACTGAGCGCTTCAACTCT-GAGATCAAGCATCTCTGCTCAGCCTCC 550
QY 180 In-ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuLeuPheIlePhe 199
DB 551 AAGTAGCTGGGACCAACAGACATGACCACTACCTGCTAATTTTATTTTATTTTATTTT 610
QY 200 AsnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeu 219
DB 611 AATTTTGTAGACAGAGTCTCAACTCTGTACCCAGGCTGGAGTGCAGTGGCGCATCTTT 670
QY 220 GlySerLeuGlnProLeuProProGlyPheLysLeuPheSerCysProSerLeuLeuSer 239
DB 671 GGCTCACTCAACCTCTGCTCCCGGGTTCAAGTTATTCTCTGCGCCAGCCTCCTCGAGT 730
QY 240 SerTrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeuValGluMet 259
DB 731 AGCTGGGATACAGGCGGCCACAGCCCTAGCTAATTTTGTATTTTATTTAGTAGAGATG 790
QY 260 -GlyPheThrMetPheAlaArgLeuLeu-LeuSerGlyProCysAspLeuProAlaAs 279
DB 791 GGGTTTACCATGTTCCGAGGTTGATGTAGATCTCTTGACCTTGTGATCTGCTGCTGCT 850
QY 279 eAlaSerGlnSerAlaGlyIleThrGlyValSerHisHis-AlaArgLeuLeuPheAsn 298
```

```
DB 851 CGGCTCTCCAAAGTGTGGGATTACAGGACGTGACGCCACCCGCCCGGCTATTTTAAAT 910
QY 299 PheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsn 318
DB 911 TTTTGTGTTGTTGAAATGGAATCTCACTCTGTTTACCAGGCTGGAGTGCATGGCCAAAT 970
QY 319 LeuGlySerLeuGlnProLeuProProGlyLeuLysArgPheSerCysLeuSerLeuPro 338
DB 971 CTCGGCTCACTCAACCTCTGCTCTCCGGGCTCAAGCGATTCTCTGCTCAGCCTCCCA 1030
QY 339 SerSerTrpAspTyrGlyHisLeuProProHisLeuProAlaAsnPheCysIlePheIleArg 358
DB 1031 AGCAGCTGGGATTACGGGCACCTGCA-CCACACCCCGCTAAATTTTGTATTTTATTTAG 1089
QY 359 GlyGlyValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375
DB 1090 GCGGGGTTTACCATAATTTGTGAGGCTGGTCTCAAACTCCTGACCTCAGG 1140
```

## RESULT 14

AR051479 LOCUS 1381 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 49 from patent US 5830670.

AR051479 ACCESSION

AR051479.1 GI:5974843

KEYWORDS

SOURCE Unknown.

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 1381)

AUTHORS de la Monte, S. and Wands, J.R.

TITLE Neural thread protein gene expression and detection of Alzheimer's

JOURNAL disease

Patent: US 5830670-A 49 03-NOV-1998;

FEATURES

source

1. .1381

/organism="unknown"

/mol\_type="unassigned DNA"

## ORIGIN

## Alignment Scores:

Pred. No.: 7.6e-104 Length: 1381  
Score: 1520.00 Matches: 344  
Percent Similarity: 90.16% Conservativeness: 4  
Best Local Similarity: 89.12% Mismatches: 25  
Query Match: 74.73% Indels: 22  
DB: 6 Gaps: 3

US-09-380-203-2 (1-375) x AR051479 (1-1381)

```
QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
DB 14 ATGGAGTTTTCGCTCTTTGTCAGGCTGGAGTGCATGGCGCAATCTCAGCTCACCGC 73
QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
DB 74 AACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCTCAGCCTCCCGAGTAGCTGGATT 133
QY 41 ThrGlyMetCysThrHisAlaArgLeuLeuTyrPhePheLeuValGluMetGluPhe 60
DB 134 ACAGGCATGTGCAC-CACGCTCGGCTAAATTTTGTATTTTATTTAGTAGAGTGGAGTTT 192
QY 61 --LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerA 80
DB 193 AACTCAATGTTGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCTCCCGCTCCGCG 252
QY 80 laSerGlnSerAlaArgTyrArgThrGly-----HisHisAlaArg-LeuCysLeuAla 97
DB 253 CTGCCCAAAGTCTCAGATT---ACAGGCATGAGCACCACCATGCCCGGCTCTGCTGGCT 309
QY 98 AsnPheCysGlyArgAsnArgValSerLeuMet-CysProSerTrpSerProGluLeuL 117
DB 310 AATTTTGTGGTAGAAACAGGGTTTCACTGATGTTTGTGCCCAAGCTGGTCTCTCCTGAGCTCAA 369
```

QY 117 sGlnSerThrCysLeuSerLeuProLysCysTrpAspTyrArgAlaValProG1 137  
 Db 370 CGAGTCCACCTGCTCCCAAGGCTGGGATACAGCGGCA-GCCGTGCTGG 428  
 QY 137 YLeuPheLeuPheLeuArgHisArgCysProThrLeuThrGlnAspGluValG1 157  
 Db 429 CCTTTTATTTATTTTAAAGACACAGGTGTACCACCTTACCAGGATGAAGTGA 488  
 QY 157 nTrpCysAspHisSerSerLeuGlnProSerThrProGluLeuLys-HisProProAlas 177  
 Db 489 GTGGTGTATCATCAGCTCATCGAGCTTCAACTCTCGAGATCAAGCAATCTCTGCT 548  
 QY 177 erAlaSerGlnValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuLeuPheIleP 197  
 Db 549 CAGCTCCCAAGTAGCTGGGACCAAGACATGACACCTACACCTTGGTA-ATTTTATTT 607  
 QY 197 heIlePheAsnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpA 217  
 Db 608 TTATTTTAAATTTTGTAGACAGAGTCTCAC-TCTGTCAACCCAGGCTGGAGTGGC 666  
 QY 217 rgAsnLeuGlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerL 237  
 Db 667 GCAATCTTGGCTCACTGCAACCTCTGCTCCGGGTCAAGTTATTTCTCTGCCCGAGC 726  
 QY 237 euLeuSerSerTrpAspTyrArgProProArgLeuAlaAsnPheValPheLeuV 257  
 Db 727 TCTCAGTAGCTGGGACTACAGCGCCACCGCTAGCTAATTTTGTATTTTGTAG 786  
 QY 257 alGluMet-GlyPheThrMetPheAlaArgLeuLeuLeuLeuSerGlyProCysAspLeu 276  
 Db 787 TAGAGATGGGTTTCAACATGTTCCGAGTTGATCTTGATCTCTTGACCTTGTGATCTG 846  
 QY 277 ProAlaSerAla-SerGlnSerAlaGlyIleThrGly-ValSerHisHisAlaArgLeuI 296  
 Db 847 CCTGCTCGGCTTACCCAAAGTGTGGGATTACAGTCTGACTCCACGC---CGGCCTA 903  
 QY 296 lePheAsnPheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnT 316  
 Db 904 TTTTAAATTTTGTGTGTAATGGAATCTACCTCTGTTACCCAGGTCGGAGTGCAAT 963  
 QY 316 rpProAsnLeuGlySerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeus 336  
 Db 964 GGCA-AATCTCGGCTACTCGCAACCTCTGCTCCGGGTC-AAGCGATTCTCTGTCTCA 1021  
 QY 336 erLeuProSerSerTrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIleP 356  
 Db 1022 GCCTCCCAAGCAGCTGGGATTACGGGAC-CTGCA-CCACACCCCGCTAATTTTGTATTT 1079  
 QY 356 heIleArgGlyGlyValSerProTyrLeuSerGlyTrp-SerGlnThrProAspLeuArg 375  
 Db 1080 TCATTAGAGCGGGTTTA--CCATATTTGTAGGCTGGGTCTCAAACTCCTGACCTCAGG 1137  
 RESULT 15  
 LOCUS AR072619 1381 bp DNA linear PAT 28-AUG-2000  
 DEFINITION Sequence 49 from patent US 5948634.  
 ACCESSION AR072619  
 VERSION AR072619.1 GI:9999383  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 UNCLASSIFIED.  
 REFERENCE 1 (bases 1 to 1381)  
 AUTHORS de la Monte, S. and Wands, J.R.  
 TITLE Neural thread protein gene expression and detection of alzheimer's disease  
 JOURNAL Patent: US 5948634-A 49 07-SEP-1999;  
 FEATURES Location/Qualifiers  
 1..1381  
 /organism="unknown"  
 /mol\_type="unassigned DNA"  
 ORIGIN

Alignment Scores:  
 Pred. No.: 7.6e-104 Length: 1381  
 Score: 1520.00 Matches: 344  
 Percent Similarity: 90.16% Conservative: 4  
 Best Local Similarity: 89.12% Mismatches: 25  
 Query Match: 74.73% Indels: 22  
 DB: 6 Gaps: 3  
 US-09-380-203-2 (1-375) x AR072619 (1-1381)  
 QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyValAlaIleSerAlaHisArg 20  
 Db 14 ATGGAGTTTTCGCTCTTGTGTCAGGCTGGAGTGAATGGCGCAATCTCAGCTCACCGC 73  
 QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40  
 Db 74 AACCTCCGCTCCCGGGTTCAAGCGAATTCCTCTGCTCAGCCTCCCCAGTAGCTGGATT 133  
 QY 41 ThrGlyMetCysThrHisAlaArgLeuLeuLeuTyrPhePheLeuValGluMetGluPhe 60  
 Db 134 ACAGCATGTGCAC-CACGCTCGGCTAATTTTGTATTTTATTTTGTAGTAGAGATGGAGTT 192  
 QY 61 --LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspSerProSerValSerA 80  
 Db 193 AACTCCATGTTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCTCCGCTCTCGGC 252  
 QY 80 laSerGlnSerAlaArgTyrArgThrGly-----HisHisAlaArg-LeuCysLeuAla 97  
 Db 253 CTGCCCAAGTGTCTGAGATT---ACAGGCATGAGCCACCATGCCCGGCTCTGCTGGCT 309  
 QY 98 AsnPheCysGlyArgAsnArgValSerLeuMet-CysProSerTrpSerProGluLeuLeu 117  
 Db 310 AATTTTGTGGTAGAAACAGGGTTTCACTGATGTTGCCCAAGCTGGTCTCTCAGAGCTCAA 369  
 QY 117 sgInSerThrCysLeuSerLeuProLysCysTrpAspTyrArgAlaValProG1 137  
 Db 370 GCAGTCCACCTGCTCAGCTCCCAAGTGTGGGATTACAGCGGTCAC-GCCGTGCTGG 428  
 QY 137 YLeuPheLeuLeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValG1 157  
 Db 429 CCTTTTATTTATTTTAAAGACACAGAGTGTACCACTTATCCAGGATGAAGTGCA 488  
 QY 157 nTrpCysAspHisSerSerLeuGlnProSerThrProGluLeuLys-HisProProAlas 177  
 Db 489 GTGGTGTGATACAGCTCATCGAGCTTCAACTCTGAGATCAAGCAATCTCTCTGCTC 548  
 QY 177 erAlaSerGlnValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuLeuPheIleP 197  
 Db 549 CAGCTCCCAAGTAGCTGGGACCAAGACATGACACCTACACCTTGGTA-ATTTTATTT 607  
 QY 197 heIlePheAsnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpA 217  
 Db 608 TTATTTTAAATTTTGTAGACAGAGTCTCAC-TCTGTCAACCCAGGCTGGAGTGGC 666  
 QY 217 rgAsnLeuGlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerL 237  
 Db 667 GCAATCTTGGCTCACTGCAACCTCTGCTCCGGGTCAAGTTATTTCTCTGCTGCCAGCC 726  
 QY 237 euLeuSerSerTrpAspTyrArgProProArgLeuAlaAsnPheValPheLeuV 257  
 Db 727 TCTCAGTAGCTGGGACTACAGCGCCACCGCTAGCTAATTTTGTATTTTGTAG 786  
 QY 257 alGluMet-GlyPheThrMetPheAlaArgLeuLeuLeuLeuSerGlyProCysAspLeu 276  
 Db 787 TAGAGATGGGTTTCAACATGTTCCGAGTTGAGTGTGATCTCTTGACCTTGTGATCTG 846  
 QY 277 ProAlaSerAla-SerGlnSerAlaGlyIleThrGly-ValSerHisHisAlaArgLeuI 296  
 Db 847 CCTGCTCGGCTTACCCAAAGTGTGGGATTACAGTCTGACTCCACGC---CGGCCTA 903  
 QY 296 lePheAsnPheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnT 316  
 Db 904 TTTTAAATTTTGTGTGTAATGGAATCTACCTCTGTTACCCAGGTCGGAGTGCAAT 963  
 QY 316 rpProAsnLeuGlySerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeus 336  
 Db 964 GGCA-AATCTCGGCTACTCGCAACCTCTGCTCCGGGTC-AAGCGATTCTCTGTCTCA 1021  
 QY 336 erLeuProSerSerTrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIleP 356  
 Db 1022 GCCTCCCAAGCAGCTGGGATTACGGGAC-CTGCA-CCACACCCCGCTAATTTTGTATTT 1079  
 QY 356 heIleArgGlyGlyValSerProTyrLeuSerGlyTrp-SerGlnThrProAspLeuArg 375  
 Db 1080 TCATTAGAGCGGGTTTA--CCATATTTGTAGGCTGGGTCTCAAACTCCTGACCTCAGG 1137

Qy	316	rpProAsnLeuGlySerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuS	336
Db	964	GGCA-AATCTCGGCTACTCGCAACCTCTGCCTCCCGGTC-AAGCGATTCTCTGTCTCA	1021
Qy	336	erLeuProSerSerTrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIleP	356
Db	1022	GCCTCCCAAGCAGCTGGGATTACGGAC-CTGCA-CCACACCCCGCTAAATTTTGTATTT	1079
Qy	356	heIleArgGlyGlyValSerProTyrLeuSerGlyTrp-SerGlnThrProAspLeuArg	375
Db	1080	TCATTAGAGCGGGTTTA--CCATATTGTGAGCTGGGTCTCAAACTCCTGACCTCAGG	1137

Search completed: September 16, 2005, 00:04:01  
 Job time : 5498 secs

THIS PAGE BLANK (USPTO)

## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: September 13, 2005, 19:25:51; Search time 65 seconds  
(without alignments)  
2954.303 Million cell updates/sec

Title: US-09-380-203-2  
Perfect score: 2034  
Sequence: 1 MEFSLLPLRLCNGAISAHR.....PIRGVSPYLSGWSQTPDLR 375

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: UniProt\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2034	100.0	375	2	O60448
2	595	29.3	587	2	ALU2_HUMAN
3	568	27.9	587	1	ALU3_HUMAN
4	549.5	27.0	591	1	ALU8_HUMAN
5	498.5	24.5	591	1	ALU5_HUMAN
6	389	19.1	585	1	ALU5_HUMAN
7	377.5	18.6	158	2	Q8NAL9
8	357.5	17.6	593	1	ALU6_HUMAN
9	356.5	17.5	603	1	ALU4_HUMAN
10	349	17.2	157	2	Q8N845
11	337.5	16.6	164	2	Q6ZTS0
12	332	16.3	176	2	Q6ZR92
13	331	16.3	170	2	Q6ZV14
14	330.5	16.2	593	1	ALU7_HUMAN
15	325	16.0	143	2	Q6ZSJ7
16	320	15.7	195	2	Q6ZUW2
17	316	15.5	152	2	Q9NX85
18	303.5	14.9	239	2	Q6ZWA9
19	303	14.9	124	2	Q6ZVF6
20	301	14.8	156	2	Q8N266
21	295	14.5	193	2	Q6ZVX8
22	294.5	14.5	148	2	Q6ZUC5
23	290	14.3	123	2	Q6ZV58
24	287.5	14.1	166	2	Q6ZUC0
25	287	14.1	180	2	Q8N7M7
26	286.5	14.1	130	2	Q6ZSV2
27	286.5	14.1	131	2	Q6ZTD3
28	283	13.9	125	2	Q6ZTE1
29	281.5	13.8	147	2	Q6ZUA2
30	280	13.8	165	2	Q6ZTF6
31	279.5	13.7	132	2	Q6ZS28

32	279	13.7	142	2	Q6ZWD5	Q6zwd5 homo sapien
33	276	13.6	129	2	Q6ZW43	Q6zw43 homo sapien
34	274	13.5	125	2	Q8N8C2	Q8n8c2 homo sapien
35	274	13.5	164	2	Q6ZUK0	Q6zuko homo sapien
36	273.5	13.4	118	2	Q9H387	Q9h387 homo sapien
37	272	13.4	140	2	Q6ZUL3	Q6zul3 homo sapien
38	271	13.3	121	2	Q6ZQS8	Q6zqs8 homo sapien
39	269	13.2	137	2	Q6ZS97	Q6zs97 homo sapien
40	269	13.2	241	2	Q6ZTY3	Q6sty3 homo sapien
41	268.5	13.2	129	2	Q6ZUW8	Q6zuw8 homo sapien
42	268	13.2	260	2	Q6ZQR8	Q6zqr8 homo sapien
43	267.5	13.2	165	2	Q6ZS53	Q6zs53 homo sapien
44	267.5	13.2	165	2	Q6ZT71	Q6zt71 homo sapien
45	267	13.1	120	2	Q6ZSU7	Q6zsu7 homo sapien

## ALIGNMENTS

RESULT 1  
O60448 PRELIMINARY; PRT; 375 AA.  
ID O60448  
AC O60448;  
DT 01-AUG-1998 (Tremblrel. 07, Created)  
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Neuronal thread protein AD7c-NTP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Neuronal;  
RX MEDLINE=98064067; PubMed=9399956;  
RA Monte S.M., Ghanbari K., Frey W.H., Beheshti I., Averbach P.,  
HAUSER S.L., Ghanbari H.A., Wands J.R.;  
RT "Characterization of the AD7C-NTP cDNA expression in Alzheimer's  
disease and measurement of a 41-kD protein in cerebrospinal fluid.";  
RL J. Clin. Invest. 100:3093-3104(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Neuronal;  
RA de la Monte S.M., Wands J.R.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF010144; AAC08737.1; -;  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR GO; GO:0016021; C:integral to membrane; TAS.  
DR GO; GO:0006915; P:apoptosis; TAS.  
DR GO; GO:0007417; P:central nervous system development; TAS.  
SQ SEQUENCE 375 AA; 41720 MW; 955443950A5BFEFD CRC64;

Query Match	100.0%; Score 2034; DB 2; Length 375;
Best Local Similarity	100.0%; Pred. No. 1.9e-171;
Matches 375; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 MEFSLLPLRLCNGAISAHRNLRLPGSSDSDSPASVPVAGITGCTHARLLILFFLVNEVF 60
Db	1 MEFSLLPLRLCNGAISAHRNLRLPGSSDSDSPASVPVAGITGCTHARLLILFFLVNEVF 60
Qy	61 LHVQAGLELPTSDPSVSASQSAARYRTGTHARLCLANFCGRNRRVSLMCPMSPELKQST 120
Db	61 LHVQAGLELPTSDPSVSASQSAARYRTGTHARLCLANFCGRNRRVSLMCPMSPELKQST 120
Qy	121 CLSLPKCWYDYPRAAVPGLFILFRLHRCPTTQDEVQWCDHSSLOPSTPEIKHPASASQ 180
Db	121 CLSLPKCWYDYPRAAVPGLFILFRLHRCPTTQDEVQWCDHSSLOPSTPEIKHPASASQ 180
Qy	191 VAGTKDMHHYTLWLFIFLNFRLROSLNSVTQAGVQWRNLGSLQPLPGKFLFSCPSLLSS 240
Db	191 VAGTKDMHHYTLWLFIFLNFRLROSLNSVTQAGVQWRNLGSLQPLPGKFLFSCPSLLSS 240
Qy	241 WDYRRPRLANFFVFLVEMGFTMFARLLILISGPCDLPASASQAGITGVSHARLIFNFC 300

```

Db      241 WYRRPRLANFVFLVEMGFTWFAKLILISGCDLPASASAGITGVSHHARLIFNFC 300
QY      301 LFEMESHVTOAGVQWPNLGSLOPLPPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGG 360
Db      301 LFEMESHVTOAGVQWPNLGSLOPLPPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGG 360
QY      361 VSPYLSGWSQTPDLR 375
Db      361 VSPYLSGWSQTPDLR 375

```

## RESULT 2

```

ALU2_HUMAN
ID ALU2_HUMAN STANDARD; PRT; 587 AA.
AC P39189;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alu subfamily SB sequence contamination warning entry.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;
RA Claverie J.-M., Makalowski W.;
RT "Alu alert.";
RL Nature 371:752-753 (1994).
RN [2]
RP CONCEPT.
RX MEDLINE=92241891; PubMed=1572661;
RA Claverie J.-M.;
RT "Identifying coding exons by similarity search: alu-derived and other
RT potentially misleading protein sequences.";
RL Genomics 12:838-841 (1992).
RN [3]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE=88333009; PubMed=3138422;
RA "The Alu family developed through successive waves of fixation closely
RA connected with primate lineage history.";
RL J. Mol. Evol. 27:194-202 (1988).
RN [4]
RP ALU FAMILIES CLASSIFICATION.
RX MEDLINE=91178815; PubMed=1706781;
RA Jurka J., Milosavljevic A.;
RT "Reconstruction and analysis of human Alu genes.";
RL J. Mol. Evol. 32:105-121 (1991).
CC -!- MISCELLANEOUS: Various analyses indicate that Alu repeats fall
CC into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning
CC consensus sequences have been constituted that contain all six
CC frames conceptual translations of each of these classes of Alu
CC repeats.
CC -!- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop
CC codon, 'XXX' is used to separate the various translation phases.
CC -!- CAUTION: This Alu entry is provided in order to avoid the further
CC pollution of protein sequence databases with Alu-derived amino
CC acid sequences.
CC -!- CAUTION: Alu repetitive sequences are interspersed in human and
CC primate genomes with an average spacing of 4 kb. Some of them are
CC actively transcribed by pol III. Normal transcripts may contain
CC Alu-derived sequences in 5' or 3' untranslated regions. However,
CC cDNA libraries also contain partial and/or rearranged cDNAs
CC ligated with Alu-derived sequence in any orientation. Although Alu
CC elements (especially situated on the complementary strand) have a
CC great potential to create additional/alternative exons,
CC consideration should be given to the possibility that the presence
CC of an Alu in an open reading frame may have resulted from a
CC cloning artifact or may be due to misinterpretation of sequencing
CC data. This point has been overlooked on several occasions, with
CC the consequence of erroneous Alu-derived amino acid sequences

```

```

CC being reported.
CC -!- CAUTION: Any significant similarity of a putative protein sequence
CC with an Alu-translated entry must be taken as a warning that a
CC part of Alu repeat may have been artifactually included in the
CC coding nucleotide sequence.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U14568; -, NOT ANNOTATED_CDS.
KW Hypothetical protein.
FT DOMAIN 1 96 Frame-1.
FT DOMAIN 100 194 Frame-2.
FT DOMAIN 198 292 Frame-3.
FT DOMAIN 296 391 Frame-4.
FT DOMAIN 395 489 Frame-5.
FT DOMAIN 493 587 Frame-6.
SQ SEQUENCE 587 AA; 63703 MW; 3EAA3E3E3929203 CRC64;
Query Match 29.3%; Score 595; DB 1; Length 587;
Best Local Similarity 41.5%; Pred. No. 5.7e-44;
Matches 153; Conservative 14; Mismatches 96; Indels 106; Gaps 8;
QY 6 LLPLECNGAISAHNRNLRLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEFHVGQ 65
Db 304 LSPRLCSGSAISAHCKLRLPGSRHSPASASRVAGTTGARHARLI-FVFLVETGFHRSQ 362
QY 66 AGLELPTSDPSSVSASQASARYTGHARLCLANFCGRNVRVSLMCPSPKQSTCLSLP 125
Db 363 DGLDLT-----SXSAR-----LGLP 378
QY 126 KCWDYRRAAVPGLFILFLHRCPLTODEVQWCHDSLSQSTPEIKHPASASQAVGK 185
Db 379 KCWDYRREPPRPAXXXFF-----XDGVSLLC-----RFGWSAVASRLTASSAS--- 421
QY 186 DMHHTYLIFIFIFNFLRQSLNSVTQAGVQWRNLGSLQPLPPGKFLFCPSLLSSWDYRR 245
Db 422 -----RVHAILLPQPEXGLQAPATTP----- 444
QY 246 PPRLANFVFLVEMGFTWFAKLILISGCDLPASASQAGITGVSHHARLIFNFCLEME 305
Db 445 ---GXFLYFXRRGFTVLARMVSI SXPRDPPASASQAGITGVSHRAXXX---FFETE 497
QY 306 SHSVTQAGVQWPNLGSLOPLPPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGVSPLY 365
Db 498 SRVAQAGVQWRDLGSLQAPPPGFTPFSCLSLPSWDYRRPPRPANFCIFSRDGVSPGX 557
QY 366 SGWSQTPDL 374
Db 558 PGWSRSPDL 566
RESULT 3
ALU3_HUMAN
ID ALU3_HUMAN STANDARD; PRT; 587 AA.
AC P39190;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alu subfamily SBI sequence contamination warning entry.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;
RA Claverie J.-M., Makalowski W.;

```

RT "Alu alert."; 6  
 RL Nature 371:752-752 (1994).  
 RN [2]  
 RP MEDLINE=92241891; PubMed=1572661;  
 RX Claverie J.-M.;  
 RA "Identifying coding exons by similarity search: alu-derived and other  
 RT potentially misleading protein sequences.";  
 RL Genomics 12:838-841 (1992).  
 RN [3]  
 RP ALU FAMILIES CLASSIFICATION.  
 RX MEDLINE=88333009; PubMed=3138422;  
 RA Quentin Y.;  
 RT "The Alu family developed through successive waves of fixation closely  
 RL connected with primate lineage history.";  
 RN J. Mol. Evol. 27:194-202 (1988).  
 RP ALU FAMILIES CLASSIFICATION.  
 RX MEDLINE=91178815; PubMed=1706781;  
 RA Jurka J., Milosavljevic A.;  
 RT "Reconstruction and analysis of human Alu genes.";  
 RL J. Mol. Evol. 32:105-121 (1991).  
 CC -!- MISCELLANEOUS: Various analyses indicate that Alu repeats fall  
 CC into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning  
 CC consensus sequences have been constituted that contain all six  
 CC frames conceptual translations of each of these classes of Alu  
 CC repeats.  
 CC -!- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop  
 CC codon, 'XXX' is used to separate the various translation phases.  
 CC -!- CAUTION: This Alu entry is provided in order to avoid the further  
 CC pollution of protein sequence databases with Alu-derived amino  
 CC acid sequences.  
 CC -!- CAUTION: Alu repetitive sequences are interspersed in human and  
 CC primate genomes with an average spacing of 4 kb. Some of them are  
 CC actively transcribed by pol III. Normal transcripts may contain  
 CC Alu-derived sequences in 5' or 3' untranslated regions. However,  
 CC cDNA libraries also contain partial and/or rearranged cDNAs  
 CC ligated with Alu-derived sequence in any orientation. Although Alu  
 CC elements (especially situated on the complementary strand) have a  
 CC great potential to create additional/alternative exons,  
 CC consideration should be given to the possibility that the presence  
 CC of an Alu in an open reading frame may have resulted from a  
 CC cloning artifact or may be due to misinterpretation of sequencing  
 CC data. This point has been overlooked on several occasions, with  
 CC the consequence of erroneous Alu-derived amino acid sequences  
 CC being reported.  
 CC -!- CAUTION: Any significant similarity of a putative protein sequence  
 CC with an Alu-translated entry must be taken as a warning that a  
 CC part of Alu repeat may have been artifactually included in the  
 CC coding nucleotide sequence.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U14569; -, NOT ANNOTATED\_CDS.  
 DR Hypothetical protein.  
 KW DOMAIN 1 96 Frame-1.  
 FT DOMAIN 100 194 Frame-2.  
 FT DOMAIN 198 292 Frame-3.  
 FT DOMAIN 296 391 Frame-4.  
 FT DOMAIN 395 489 Frame-5.  
 FT DOMAIN 493 587 Frame-6.  
 SQ SEQUENCE 587 AA; 63573 MW; 85C4155726DEF235 CRC64;  
 Query Match 27.9%; Score 568; DB 1; Length 587;  
 Best Local Similarity 41.5%; Pred. No. 1.4e-41;  
 Matches 153; Conservative 15; Mismatches 95; Indels 106; Gaps 10;

QY 6 LLPLRLCNGAISAHNNLRPLPGSSDSPASASPVAGITGMCTHARLILYFFLVEMFLRVGQ 65  
 DB LSPRLCEGSGISAHCKLRPLPGSRHSPASASQVAGITGARHYARLI-FVELVETGFFRFSR 362  
 QY 66 AGLELPTSDPVSASQASARYFTGHARILCLANFCGRNVRVSLMCPSPASQVAGTK 125  
 DB DGLDLTLT-----SXSAR-----XDGVSLLC-----RCWSA--VAGSR 414  
 QY 126 KCMYDERRAIVPGLFILFLRHRCPRTLQDEVOVQWCHSSLOPSTPEIKHPASASQVAGTK 185  
 DB DGLDLTLT-----SXSAR-----XDGVSLLC-----RCWSA--VAGSR 414  
 QY 186 DMHHTYTLWIFIFINFLRQSLNSVTQAGVQWNLGSLQPLPFGKLFSCPSLLSSWDYRR 245  
 DB -----LTASSASRVHA-----ILLPQPKXGLQAPATTP----- 444  
 QY 246 PPRLANFFVFLVEMGFTMPARILILIGPCDLSPASASQASAGITGVSHHAKLIINFCLFEME 305  
 DB -----GXFLYFXRRRGFTVLGAMVSVISXPRDPPASASQASAGITGVSHHAKLXXX--PFETE 497  
 QY 306 SHSVTQAGVQWNLGSLQPLPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGGVSPYL 365  
 DB SRSVQAQVQWRDLGSLQAPPGFTFFSCLSLPSWDYRRPRLPANFCIFSRDGVSPFX 557  
 QY 366 SCWSQTPDL 374  
 DB PMSRSRSPDL 566  
 RESULT 4  
 ID ALUB HUMAN STANDARD; PRT; 591 AA.  
 AC P39195;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Alu subfamily SX sequence contamination warning entry.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;  
 RA Claverie J.-M., Makalowski W.;  
 RT "Alu alert.";  
 RL Nature 371:752-752 (1994).  
 RN [2]  
 RP CONCEPT.  
 RX MEDLINE=92241891; PubMed=1572661;  
 RA Claverie J.-M.;  
 RT "Identifying coding exons by similarity search: alu-derived and other  
 RL potentially misleading protein sequences.";  
 RL Genomics 12:838-841 (1992).  
 RN [3]  
 RP ALU FAMILIES CLASSIFICATION.  
 RX MEDLINE=88333009; PubMed=3138422;  
 RA Quentin Y.;  
 RT "The Alu family developed through successive waves of fixation closely  
 RL connected with primate lineage history.";  
 RL J. Mol. Evol. 27:194-202 (1988).  
 RN [4]  
 RP ALU FAMILIES CLASSIFICATION  
 RX MEDLINE=91178815; PubMed=1706781;  
 RA Jurka J., Milosavljevic A.;  
 RT "Reconstruction and analysis of human Alu genes.";  
 RL J. Mol. Evol. 32:105-121 (1991).  
 CC -!- MISCELLANEOUS: Various analyses indicate that Alu repeats fall  
 CC into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning  
 CC consensus sequences have been constituted that contain all six  
 CC frames conceptual translations of each of these classes of Alu  
 CC repeats.  
 CC -!- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop

```
CC codon, 'XXX' is used to separate the various translation phases.
CC -! CAUTION: This Alu entry is provided in order to avoid the further
CC pollution of protein sequence databases with Alu-derived amino
CC acid sequences.
CC -! CAUTION: Alu repetitive sequences are interspersed in human and
CC primate genomes with an average spacing of 4 kb. Some of them are
CC actively transcribed by pol III. Normal transcripts may contain
CC Alu-derived sequences in 5' or 3' untranslated regions. However,
CC cDNA libraries also contain partial and/or rearranged cDNAs
CC ligated with Alu-derived sequence in any orientation. Although Alu
CC elements (especially situated on the complementary strand) have a
CC great potential to create additional/alternative exons,
CC consideration should be given to the possibility that the presence
CC of an Alu in an open reading frame may have resulted from a
CC cloning artifact or may be due to misinterpretation of sequencing
CC data. This point has been overlooked on several occasions, with
CC the consequence of erroneous Alu-derived amino acid sequences
CC being reported.
CC -! CAUTION: Any significant similarity of a putative protein sequence
CC with an Alu-translated entry must be taken as a warning that a
CC part of Alu repeat may have been artifactually included in the
CC coding nucleotide sequence.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U14574; -; NOT_ANNOTATED_CDS.
CC KW Hypothetical protein.
CC FT DOMAIN 1 96 Frame-1.
CC FT DOMAIN 100 195 Frame-2.
CC FT DOMAIN 199 294 Frame-3.
CC FT DOMAIN 298 393 Frame-4.
CC FT DOMAIN 397 492 Frame-5.
CC FT DOMAIN 496 591 Frame-6.
CC SQ SEQUENCE 591 AA; 64395 MW; AC8154AD8A6B280 CRC64;
CC -----
CC Query Match 27.0%; Score 549.5; DB 1; Length 591;
CC Best Local Similarity 38.9%; Pred. No. 6.1e-40;
CC Matches 150; Conservative 15; Mismatches 84; Indels 137; Gaps 11;
CC -----
CC Qy* 6 LPRLECGAISAHNLRLPGSSDSPASASPVAGITGCMTHARLILYFFLVEMEFHVGQ 65
CC Db 306 LSPRLCSCAISAHCNLRPLPGSSDSPASASRVAGITGARHRLI-FVFLVETGFHHVQ 364
CC -----
CC Qy 66 AGLELPTSD--PSVSASQARYRTCHHARLCLANFCGRNRYSLMCPWS-----P 114
CC Db 365 AGLELTTGDLPPKVLGLQAXATAPQ---LRPXXXFFDXGVSCLRPGMSAVARSRLTAT 420
CC Qy 115 ELKQSTCLSLPKWD---YRAAVPGLFILFLRHRCPTLTQDEVQWCHDSLSLOPSPE 170
CC Db 421 SASRVQAILLPQPEXGLQARATTPCXFLYFXRRGFTMLAR-----LVNSXPQ 471
CC Qy 171 IKHPPASQVAGTKDMHHYTLWIFIFNLFRLQSLNSVTQAGVQWRNLGSLQPLPGPK 230
CC Db 472 V-----IF----- 474
CC Qy 231 LFSCLSLSSWDYRR-PPRLANFFVFLVEMGTFMFAILLISGPCDLPASASQSGAGITGV 289
CC Db 475 -----LPKCWDYRREPP-----PASARXXX----- 495
CC Qy 290 SHHARLIFNFCFLFEMESHVSTQAGVQWRNLGSLQPLPGKRFKCSLSPSSWDYGHLP 349
CC Db 496 -----PFETESRVAQAGVQWRDLGSLQPPPGFKFCSLSPSSWDYRRAPPR 544
CC -----
CC 350 PANFCIFIRGGVSPYLSGWSQTPDLR 375
CC Db 545 PANFCIFIRSDGVSPCPGWSRTPDLR 570
```

```
RESULT 5
ALU1_HUMAN STANDARD; PRT; 591 AA.
AC P39188;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 25-OCT-2004 (Rel. 45, Last annotation update)
DE Alu subfamily J sequence contamination warning entry.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;
RA Claverie J.-M., Makalowski W.;
RT "Alu alert.";
RL Nature 371:752-752(1994).
RN [2]
CONCEPT.
RX MEDLINE=92241891; PubMed=1572661;
RA Claverie J.-M.;
RT "Identifying coding exons by similarity search: alu-derived and other
RT potentially misleading protein sequences.";
RL Genomics 12:838-841(1992).
RN [3]
ALU FAMILIES CLASSIFICATION.
RX MEDLINE=89333009; PubMed=3138422;
RA Quentin Y.;
RT "The Alu family developed through successive waves of fixation closely
RT connected with primate lineage history.";
RL J. Mol. Evol. 27:194-202(1988).
RN [4]
ALU FAMILIES CLASSIFICATION.
RX MEDLINE=91178815; PubMed=1706781;
RA Jurka J., Milosavljevic A.;
RT "Reconstruction and analysis of human Alu genes.";
RL J. Mol. Evol. 32:105-121(1991).
CC -! MISCELLANEOUS: Various analyses indicate that Alu repeats fall
CC into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning
CC consensus sequences have been constituted that contain all six
CC frames conceptual translations of each of these classes of Alu
CC repeats.
CC -! MISCELLANEOUS: Isolated 'X' indicates the presence of a stop
CC codon, 'XXX' is used to separate the various translation phases.
CC -! CAUTION: This Alu entry is provided in order to avoid the further
CC pollution of protein sequence databases with Alu-derived amino
CC acid sequences.
CC -! CAUTION: Alu repetitive sequences are interspersed in human and
CC primate genomes with an average spacing of 4 kb. Some of them are
CC actively transcribed by pol III. Normal transcripts may contain
CC Alu-derived sequences in 5' or 3' untranslated regions. However,
CC cDNA libraries also contain partial and/or rearranged cDNAs
CC ligated with Alu-derived sequence in any orientation. Although Alu
CC elements (especially situated on the complementary strand) have a
CC great potential to create additional/alternative exons,
CC consideration should be given to the possibility that the presence
CC of an Alu in an open reading frame may have resulted from a
CC cloning artifact or may be due to misinterpretation of sequencing
CC data. This point has been overlooked on several occasions, with
CC the consequence of erroneous Alu-derived amino acid sequences
CC being reported.
CC -! CAUTION: Any significant similarity of a putative protein sequence
CC with an Alu-translated entry must be taken as a warning that a
CC part of Alu repeat may have been artifactually included in the
CC coding nucleotide sequence.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U14567; -; NOT\_ANNOTATED\_CDS.

CC Hypothetical protein.

CC FT DOMAIN 1 96 Frame-1.

CC FT DOMAIN 100 195 Frame-2.

CC FT DOMAIN 199 294 Frame-3.

CC FT DOMAIN 298 393 Frame-4.

CC FT DOMAIN 397 492 Frame-5.

CC FT DOMAIN 496 591 Frame-6.

CC SEQUENCE 591 AA; 63790 MW; 665D395735519D95 CRC64;

Query Match 24.5%; Score 498.5; DB 1; Length 591;

Best Local Similarity 38.6%; Pred. No. 2e-35;

Matches 144; Conservative 15; Mismatches 103; Indels 111; Gaps 11;

QY 6 LPLRLCNGAISHNRNLRPGSSDPSASAPVAGTGMCTHARLILYFFLVMEFLHVGQ 65

DB 306 LSPRLCNGAISHNRNLRPGSSDPSASAPVAGTGMCTHARLILYFFLVMEFLHVGQ 364

QY 66 AGLELPTSDPSVSQASRYRTG--HHARLCLANFCGRNRYSLMCPSPKSTCLLS 123

DB 365 AGLELPGSSDPSASQAGI-TGVSHRR--XXXFFXDRVSLCRPGWSA----- 411

QY 124 LPKCWDYRAAVPGLFILFRLHRCPTLTQDEVQWCDHSLQSTPEIKHPASASQVAG 183

DB 412 -----VARSRLTAASRAQ-----ALLPOPE----- 435

QY 184 TKDHHYTWLFIPIFNRLQSLNSTVQAGVQWRNLSLQPLPPGKFLSCPSLLSSWDY 243

DB 436 -----XLGLQAR-----ATTPG----- 447

QY 244 RPPRLANFVFLVGVGTFMFLRLILIS-GPCDLPASASQAGITGVSHHARLIFNFCFLF 302

DB 448 -----XFLYFLXRRGFTMLRLVSNWAQVILPPRPK---VLGLQAXATAPGXXXPF 497

QY 303 EMESHVTVQAGVQWPNLGLQPLPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGGVS 362

DB 498 ETGSRVQAGVQWRDHSGLQPLPGLKRSCLSLPSWDYRRAPRPNFCIFCDGVIS 557

QY 363 PYLGSWQTPDLR 375

DB 558 LCCPGWSTPGLK 570

## RESULT 6

ALUS\_HUMAN STANDARD; PRT; 585 AA.

AC P39192;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Alu subfamily SC sequence contamination warning entry.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI\_TaxID=9606;

RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;

RA Claverie J.-M., Makalowski W.;

RT "Alu alert".

RL Nature 371:752-752(1994).

RN [2]

CONCEPT.

RA MEDLINE=92241891; PubMed=1572661;

RX Claverie J.-M.;

RT "Identifying coding exons by similarity search: alu-derived and other

RT potentially misleading protein sequences".

RL Genomics 12:838-841(1992).

RN [3]

ALU FAMILIES CLASSIFICATION.

RX MEDLINE=88333009; PubMed=3138422;

RA Quentin Y.;

RT "The Alu family developed through successive waves of fixation closely

RT connected with primate lineage history".

RL J. Mol. Evol. 27:194-202(1988).

RN [4]

ALU FAMILIES CLASSIFICATION.

RX MEDLINE=91178815; PubMed=1706781;

RA Jurka J., Milosavljevic A.;

RT "Reconstruction and analysis of human Alu genes".

RL J. Mol. Evol. 32:105-121(1991).

CC -!- MISCELLANEOUS: Various analyses indicate that Alu repeats fall

CC into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning

CC consensus sequences have been constituted that contain all six

CC frames conceptual translations of each of these classes of Alu

CC repeats.

CC -!- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop

CC codon, 'XXX' is used to separate the various translation phases.

CC -!- CAUTION: This Alu entry is provided in order to avoid the further

CC pollution of protein sequence databases with Alu-derived amino

CC acid sequences.

CC -!- CAUTION: Alu repetitive sequences are interspersed in human and

CC primate genomes with an average spacing of 4 Kb. Some of them are

CC actively transcribed by pol III. Normal transcripts may contain

CC Alu-derived sequences in 5' or 3' untranslated regions. However,

CC cDNA libraries also contain partial and/or rearranged cDNAs

CC ligated with Alu-derived sequence in any orientation. Although Alu

CC elements (especially situated on the complementary strand) have a

CC great potential to create additional/alternative exons,

CC consideration should be given to the possibility that the presence

CC of an Alu in an open reading frame may have resulted from a

CC cloning artifact or may be due to misinterpretation of sequencing

CC data. This point has been overlooked on several occasions, with

CC the consequence of erroneous Alu-derived amino acid sequences

CC being reported.

CC -!- CAUTION: Any significant similarity of a putative protein sequence

CC with an Alu-translated entry must be taken as a warning that a

CC part of Alu repeat may have been artifactually included in the

CC coding nucleotide sequence.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; U14571; -; NOT\_ANNOTATED\_CDS.

DR Hypothetical protein.

KW DOMAIN 1 95 Frame-1.

FT DOMAIN 99 193 Frame-2.

FT DOMAIN 197 291 Frame-3.

FT DOMAIN 295 389 Frame-4.

FT DOMAIN 393 487 Frame-5.

FT DOMAIN 491 585 Frame-6.

SQ SEQUENCE 585 AA; 63957 MW; 46E8C4F493650A7 CRC64;

Query Match 19.1%; Score 389; DB 1; Length 585;

Best Local Similarity 33.9%; Pred. No. 9.5e-26;

Matches 130; Conservative 23; Mismatches 100; Indels 130; Gaps 17;

QY 17 SAHRNLRPLPGSSDPSASAPVAGITGMCTHARLILYFFLVMEFLHVGAGLELPTSDDP 76

DB 189 SVSKKXXPGAVAHACNPNSTLGGGRITRSL-----DRHPGGHG-----ETP 232

QY 77 SVSASQASRYRTGHHARLCLANFCGR-----NRV-----SLMC-PSWSPELKQST 120

DB 233 SLLKTKLQAGRG--ARL-XSOLLGLRQENRNLPGGGCSEPRSRHCTPAW-----RQSE 285

QY 121 CLSLPKCWDYRAAVPGLFILFRLHRCPTLTQDEVQWCDHSLQSTPEIKHPASASQ 180

DB 286 TFSQKK-----XXXFFLRSL-----ALSPGW-----SA 309



CC being reported.  
 CC -!- CAUTION: Any significant similarity of a putative protein sequence  
 CC with an Alu-translated entry must be taken as a warning that a  
 CC part of Alu repeat may have been artifactually included in the  
 CC coding nucleotide sequence.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U14572; -; NOT ANNOTATED\_CDS.  
 CC ProDom; PD003738; GIDA; 1.  
 KW Hypothetical protein.  
 FT DOMAIN 1 97 Frame-1.  
 FT DOMAIN 101 196 Frame-2.  
 FT DOMAIN 200 295 Frame-3.  
 FT DOMAIN 299 395 Frame-4.  
 FT DOMAIN 399 494 Frame-5.  
 FT DOMAIN 496 593 Frame-6.  
 SQ SEQUENCE 593 AA; 64603 MW; 136EF344AACD12A2 CRC64;  
 Query Match 17.6%; Score 357.5; DB 1; Length 593;  
 Best Local Similarity 48.2%; Pred. No. 5.9e-23;  
 Matches 95; Conservative 6; Mismatches 47; Indels 49; Gaps 8;  
 QY 4 SLLPLRLCNGAISAHNRLLPGSSDSPASVAGITGCTHARLLYFLVMEFLHV 63  
 DB 405 SLLPLRLCNGAISAHNRLLPGSSDSPASVAGITGCTHARLLYFLVMEFLHV 463  
 QY 64 GAGLELPTSDPSVSASQASRYTG--HARLCLANFCGRNRVSLMCPDWS- 113  
 DB 464 GAGLELPTSDPSVSASQASAGI-TGVSHRAR---XXFFETFRSCCPGWSAMARSRLT 519  
 QY 114 -----PELQSTCLSLPKCWDYRRAAVPGLFILFLFHRCPRTLTDDEVQW 158  
 DB 520 ATTSASRVQAILLPQPEX-----LGLQAC-----ATTPGXFLYXXRRGFSMLVR----- 564  
 QY 159 CDHSSLQSTPPIKHP 175  
 DB 565 ----LVNSRQVIRPP 577  
 RESULT 9  
 ID ALU4 HUMAN STANDARD; PRT; 603 AA.  
 AC P39191;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Alu subfamily SB2 sequence contamination warning entry.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=95021758; PubMed=7935834; DOI=10.1038/371752a0;  
 RX Claverie J.-M., Makalowski W.;  
 RT "Alu alert";  
 RL Nature 371:752-752(1994).  
 RN [2]  
 RN CONCEPT.  
 RP MEDLINE=92241891; PubMed=1572661;  
 RX Claverie J.-M.;  
 RT "Identifying coding exons by similarity search: alu-derived and other  
 RT potentially misleading protein sequences";  
 RL Genomics 12:838-841(1992).  
 RN [3]  
 RP ALU FAMILIES CLASSIFICATION.

RX MEDLINE=88333009; PubMed=3138422;  
 RA Quentin Y.;  
 RT "The Alu family developed through successive waves of fixation closely  
 RT connected with primate lineage history";  
 RL J. Mol. Evol. 27:194-202(1988).  
 RN [4]  
 RP ALU FAMILIES CLASSIFICATION.  
 RX MEDLINE=91178815; PubMed=1706781;  
 RA Jurka J., Milosavljevic A.;  
 RT "Reconstruction and analysis of human Alu genes";  
 RL J. Mol. Evol. 32:105-121(1991).  
 CC -!- MISCELLANEOUS: Various analyses indicate that Alu repeats fall  
 CC into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning  
 CC consensus sequences have been constituted that contain all six  
 CC frames conceptual translations of each of these classes of Alu  
 CC repeats.  
 CC -!- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop  
 CC codon, 'XXX' is used to separate the various translation phases.  
 CC -!- CAUTION: This Alu entry is provided in order to avoid the further  
 CC pollution of protein sequence databases with Alu-derived amino  
 CC acid sequences.  
 CC -!- CAUTION: Alu repetitive sequences are interspersed in human and  
 CC primate genomes with an average spacing of 4 kb. Some of them are  
 CC actively transcribed by pol III. Normal transcripts may contain  
 CC Alu-derived sequences in 5' or 3' untranslated regions. However,  
 CC cDNA libraries also contain partial and/or rearranged cDNAs  
 CC ligated with Alu-derived sequence in any orientation. Although Alu  
 CC elements (especially situated on the complementary strand) have a  
 CC great potential to create additional/alternative exons,  
 CC consideration should be given to the possibility that the presence  
 CC of an Alu in an open reading frame may have resulted from a  
 CC cloning artifact or may be due to misinterpretation of sequencing  
 CC data. This point has been overlooked on several occasions, with  
 CC the consequence of erroneous Alu-derived amino acid sequences  
 CC being reported.  
 CC -!- CAUTION: Any significant similarity of a putative protein sequence  
 CC with an Alu-translated entry must be taken as a warning that a  
 CC part of Alu repeat may have been artifactually included in the  
 CC coding nucleotide sequence.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U14570; -; NOT ANNOTATED\_CDS.  
 KW Hypothetical protein.  
 FT DOMAIN 1 98 Frame-1.  
 FT DOMAIN 102 199 Frame-2.  
 FT DOMAIN 202 300 Frame-3.  
 FT DOMAIN 304 401 Frame-4.  
 FT DOMAIN 405 502 Frame-5.  
 FT DOMAIN 506 603 Frame-6.  
 SQ SEQUENCE 603 AA; 65272 MW; B8AAD0AD46BEA114 CRC64;  
 Query Match 17.5%; Score 356.5; DB 1; Length 603;  
 Best Local Similarity 30.3%; Pred. No. 7.3e-23;  
 Matches 112; Conservative 28; Mismatches 115; Indels 115; Gaps 14;  
 QY 76 PVSASQASRYRTGHARLCLANFCGRNRVSLMCPDWSPELQSTCLSLPKCWDYRRAAV 135  
 DB 156 PSYWGXXRRMAXTREAEALVSRDCA---TAVRSPAWATERDSVS-----KQXXX 202  
 QY 136 PGLFILFLFHRCPRTLQDEVQWC-----DHSSLQSTP----- 169  
 DB 203 PGA-----VANACNPSTLGGCGWIMRSGDRDHPGXQGETPSLLKIQISRARWAPVVP 257  
 QY 170 -----EIKHP-----PASASQVAGTKMDHHYTWLIFIFINFLR 203  
 DB 258 ATGEAEAGEWREPGRKRSIQXAEIAPLQSAVRPGRQSTPQSKKXKF-----FLR 307



QY 204 QS--LNSVTQAGVQWNLGSLQPLPGFKLFSCPSLLSSWDYRR--PPRLANFFVFLVEMG 260  
 Db 308 RSLALSPRPDCGLQWNLGSLQAPLPGFTFFSCLSLPSSWDYRRPPRPANFLYFXRRG 367  
 QY 261 FTMFARLIILISGPCDLPASASAGITGVSHARLIFNF-----CLFEMES 306  
 Db 368 FTLLARWVISAPHPDPPASASAGITGVSHARXXXFXDGVSLCRPORTADCSAISA 427  
 QY 307 HSVTQAGVQWNLGSLQPLPGFKLFSCPSLLSSWDYHGLPHPPAN--FCIFIRGVSPY 364  
 Db 428 HCKLRP-----PCSRHSPASASPVAGTTG--ARRHARLIFCIFSROGVSPC 471  
 QY 365 LSGWSQTPDL 374  
 Db 472 XFGWSRSPDL 481

RESULT 10  
 Q8N845 PRELIMINARY; PRT; 157 AA.  
 ID Q8N845  
 AC Q8N845;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DE Hypothetical protein FLJ40023.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Stomach;  
 RX PubMed=14702039; DOI=10.1038/ng1285;  
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita T., Ishii S.,  
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,  
 RA Murakami K., Tasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,  
 RA Sudo H., Hosiiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,  
 RA Abe K., Kamihara K., Katsuma N., Sato K., Tanikawa M., Yamazaki M.,  
 RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,  
 RA Tani H., Kimata M., Watanabe M., Hirao K., Chiba Y., Ishida S.,  
 RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,  
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,  
 RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,  
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,  
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Okamoto S.,  
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
 RA Matsunaga K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
 RA Togashi T., Oyama H., Hata H., Watanabe M., Komatsu T.,  
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isoigai T., Sugano S.;  
 RT "Complete sequencing and characterization of 21,243 full-length human  
 cDNAs";  
 RL Nat. Genet. 36:40-45(2004).  
 DR EMBL; AK097342; BAC05007.1; -;  
 SQ SEQUENCE 157 AA; 17930 MW; D79F77F3B7C608E CRC64;

Query Match 17.2%; Score 349; DB 2; Length 157;  
 Best Local Similarity 45.9%; Pred. No. 7.3e-23;  
 Matches 85; Conservative 15; Mismatches 47; Indels 38; Gaps 6;  
 QY 189 HTYTLFIIFIFNLRQSLNSVTQAGVQWNLGSLQPLPGFKLFSCPSLLSSWDYRRPPR 248

Db 5 HHANLIFLFL---VEMGFHVQASVQ-----LLTSKQL----- 35  
 QY 249 LANFFVFLVEMGFTMFA--RLILISGPCDLPASASQSGAGITGVSHH---RLIFNCLFE 303  
 Db 36 ---IFVFLIETGFRHVAQPRLELL--GSNDPPTSASQAGIIGMSHHTWPFQFFYFFPFE 91  
 QY 304 MESHVVTQAGVQWNLGSLQPLPGFKLFSCPSLLSSWDYHGLPHPPANFCIFIRGVSP 363  
 Db 92 TESRSVAQAGVQWRDLGSLKPPSPGFKQFQSCLSLLSSWDYRCLPCLANFCIFSRDRLP 151  
 QY 364 YLSGW 368  
 Db 152 HCPGW 156

RESULT 11  
 Q6ZTSO PRELIMINARY; PRT; 164 AA.  
 ID Q6ZTSO  
 AC Q6ZTSO;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DE Hypothetical protein FLJ44280.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Trachea;  
 RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,  
 RA Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,  
 RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,  
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,  
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,  
 RA Wagatsuma M., Murakawa K., Kanehori K., Sugiyama A., Kawakami B.,  
 RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isoigai T.,  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK126268; BAC86510.1; -;  
 SQ SEQUENCE 164 AA; 17866 MW; 7290204DF0901991 CRC64;  
 Query Match 16.6%; Score 337.5; DB 2; Length 164;  
 Best Local Similarity 36.4%; Pred. No. 8e-22; Indels 115; Gaps 6;  
 Matches 95; Conservative 8; Mismatches 43;  
 QY 43 MCHARLILYFFLVEMEFHVQAGLELPTSDPSVSASQSAARYTGHARLCLANFCGR 102  
 Db 1 MCHARLILYFFLVETGFLHVQAGLELLTSGD----- 33  
 QY 103 NRVSIMCPSWSPKQSTCLSLPKCWYRRRAAVPGLFILFLRHRCPTLTQDEVQWCDHS 162  
 Db 34 ----- 33  
 QY 163 SLQSTPTEIKHPPASASOVAGTKDMHH-----YTWLIFIFNLRQSLNSVTQAGVQ 215  
 Db 34 -----PPASASAGITGMSHCAQPTISIFHVFLCVIFFSRWSL-SVAQAGVQ 81  
 QY 216 WRNLGSLQPLPGFKLFSCPSLLSSWDYRRPRLANFFVFLVEMGFTWFARL---ILISG 272  
 Db 82 WRDLGSL---PGFKRFSCLSLPSNWDCHRPSPCAKCTCFVMEFHHVQAGLELLTSG 137  
 QY 273 PCDLIPASASQSGAGITGVSHHA 293  
 Db 138 --DLPTLASQSGAGITGVSHHA 156

RESULT 12  
 Q6ZR92 PRELIMINARY; PRT; 176 AA.  
 ID Q6ZR92  
 AC Q6ZR92;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)



DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein FLJ46542.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Thymus;  
 RA Ninomiya K., Magatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,  
 RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kanihara K.,  
 RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,  
 RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,  
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,  
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H.,  
 RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,  
 RA Wakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,  
 RA Isogai T.,  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK128399; BAC87422.1; -  
 SQ SEQUENCE 176 AA; 19060 MW; 9BD1CAC1940AD8FA CRC64;  
 Query Match 16.3%; Score 332; DB 2; Length 176;  
 Best Local Similarity 47.9%; Pred. No. 2.7e-21;  
 Matches 90; Conservative 11; Mismatches 51; Indels 36; Gaps 8;  
 QY 208 SVTQAGVQWNLGSLQPLPPGKLFSCPSLLSSWDYRRPRLAN-----PF 253  
 DB 5 SVSQAGVQWRD-----PPAHCIL---HLSSSD---SPASASRVAVISGTCCHAQLIF 51  
 QY 254 VFLVEMGFTMFA-----RLILISGCDLPASASQAGITGVSHARLIFNF-CLFEMESH 308  
 DB 52 VFLVETGFCVVGQAGRLRLTSSDP---PTSASLSAGIAGLSHRAPFCFLCFVETVSHS 108  
 QY 309 VTQAGVQWNLGSLQPLPPGKLFSCPSLLSSWDYCHLPPHPAN-FCIFRGVSPYLSG 367  
 DB 109 VTQNGQWCHDSQSRPPGLKRSPLSLSSWDYRQASPHLANFYFFIRGKSLCCPG 168  
 QY 368 WSQTDLR 375  
 DB 169 WSQTPELK 176  
 PRELIMINARY; PRT; 170 AA.  
 AC Q6ZV14  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DE Hypothetical protein FLJ42545.  
 OS Homo sapiens (Human)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Cerebellum;  
 RA Ninomiya K., Magatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,  
 RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kanihara K.,  
 RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,  
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,  
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,  
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,  
 RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,  
 RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.,  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK124536; BAC85877.1; -  
 SQ SEQUENCE 170 AA; 19351 MW; A019980274CEA078 CRC64;  
 Query Match 16.3%; Score 331; DB 2; Length 170;  
 Best Local Similarity 42.0%; Pred. No. 3.1e-21;  
 Matches 71; Conservative 16; Mismatches 38; Indels 44; Gaps 3;

QY 251 NFFVFLVEMGFTMFAEL-----ILISGCDLPASASQAGITG-----VSH 291  
 DB 2 NLLLFKSKQGLALLPRKCSGAIHAHCNFFELGSSNFPDLASERAGTTAFGTWVLRISN 61  
 QY 292 HARLIFNF-----CLFEMESHVSTQAGVQWNLGSLQPLP 326  
 DB 62 HIAMLWDFWRRKQTIWSTRTNLNHHLVSCISFIIFETESHVSTQAGVQWNLSSLPQPP 121  
 QY 327 PGLKEPSCLSLPSSWDYCHLPPHPANFCIFRGVSPYLSGWSQTPDLR 375  
 DB 122 PGKRFPSCLTLPSTWDYQWMPCLANFCIFRSRDGVSYPWPGWSRTPDCR 170  
 RESULT 14  
 ALU7\_HUMAN STANDARD; PRT; 593 AA.  
 ID ALU7\_HUMAN  
 AC P39194;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Alu subfamily SQ sequence contamination warning entry.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=950211758; PubMed=7935834; DOI=10.1038/371752a0;  
 RA Claverie J.-M.; Makalowski W.;  
 RL "Alu alert."  
 RL Nature 371:752-752 (1994).  
 RN [2]  
 RP CONCEPT.  
 RX MEDLINE=92241891; PubMed=1572661;  
 RA Claverie J.-M.;  
 RT "Identifying coding exons by similarity search: alu-derived and other  
 RT potentially misleading protein sequences."  
 RL Genomics 12:838-841 (1992).  
 RN [3]  
 RP ALU FAMILIES CLASSIFICATION.  
 RX MEDLINE=88333009; PubMed=3138422;  
 RA Quentin Y.;  
 RT "The Alu family developed through successive waves of fixation closely  
 RT connected with primate lineage history."  
 RL J. Mol. Evol. 27:194-202 (1988).  
 RN [4]  
 RP ALU FAMILIES CLASSIFICATION.  
 RX MEDLINE=91178815; PubMed=1706781;  
 RA Jurka J., Milosavljevic A.;  
 RT "Reconstruction and analysis of human Alu genes."  
 RL J. Mol. Evol. 32:105-121 (1991).  
 CC -!- MISCELLANEOUS: Various analyses indicate that Alu repeats fall  
 CC into 8 subfamilies (Ref.3 and Ref.4). Therefore, 8 Alu warning  
 CC consensus sequences have been constituted that contain all six  
 CC frames conceptual translations of each of these classes of Alu  
 CC repeats.  
 CC -!- MISCELLANEOUS: Isolated 'X' indicates the presence of a stop  
 CC codon, 'XXX' is used to separate the various translation phases.  
 CC -!- CAUTION: This Alu entry is provided in order to avoid the further  
 CC pollution of protein sequence databases with Alu-derived amino  
 CC acid sequences.  
 CC -!- CAUTION: Alu repetitive sequences are interspersed in human and  
 CC primate genomes with an average spacing of 4 kb. Some of them are  
 CC actively transcribed by pol III. Normal transcripts may contain  
 CC Alu-derived sequences in 5' or 3' untranslated regions. However,  
 CC cDNA libraries also contain partial and/or rearranged cDNAs  
 CC ligated with Alu-derived sequence in any orientation. Although Alu  
 CC elements (especially situated on the complementary strand) have a  
 CC great potential to create additional/alternative exons,  
 CC consideration should be given to the possibility that the presence  
 CC of an Alu in an open reading frame may have resulted from a  
 CC cloning artifact or may be due to misinterpretation of sequencing

```
CC data. This point has been overlooked on several occasions, with
CC the consequence of erroneous Alu-derived amino acid sequences
CC being reported.
CC -I- CAUTION: Any significant similarity of a putative protein sequence
CC with an Alu-translated entry must be taken as a warning that a
CC part of Alu repeat may have been artifactually included in the
CC coding nucleotide sequence.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U14573; -; NOT ANNOTATED_CDS.
CC ProDom; PD003738; GIDA; 1.
CC KW Hypothetical protein.
CC FT DOMAIN 1 97 Frame-1.
CC 101 196 Frame-2.
CC FT DOMAIN 200 295 Frame-3.
CC FT DOMAIN 299 395 Frame-4.
CC FT DOMAIN 399 494 Frame-5.
CC FT DOMAIN 498 593 Frame-6.
CC SEQUENCE 593 AA; 64417 MW; 54A4F50F33A6089F CRC64;
CC -----
Query Match 16.2%; Score 330.5; DB 1; Length 593;
Best Local Similarity 45.9%; Pred. No. 1.4e-20;
Matches 89; Conservative 9; Mismatches 53; Indels 43; Gaps 7;
QY 4 SLLPLRLCNGAISAHNRRLPGSSDSPASASPVAGITGMCTHARLLILYFFLVEMEFLLHV 63
Db 405 SLLPLRLCNGAISAHNRRLPGSSDSPASASPVAGITGMCTHARLLILYFFLVEMEFLLHV 463
QY 64 GQAGLELPTSDPSVSASQSAARYTG--HHARLCLANFCGRNRVSLMCPSSWS----- 113
Db 464 GQAGLELPTSDPSVSASQSAARYTG--HHARLCLANFCGRNRVSLMCPSSWS----- 113
QY 114 -----PELQSTQCLSLPKCKDYRRRAAVPGLFILFLRHRCPTLTQDEVQWCDH 161
Db 520 ATASRVQAILLPQPEXGLQAP-----ATTGCFLYEXXRGFTMLAR----- 564
QY 162 SSLQPTPEIKHPP 175
Db 565 -LVSNSXPQVIHPP 577
CC -----
RESULT 15
Q6ZSJ7 PRELIMINARY; PRT; 143 AA.
AC Q6ZSJ7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ45460.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Subthalamic nucleus;
RA Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Suzuki Y.,
RA Sugano S., Nagahari K., Masuno Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK127382; BAC86951.1; -.
SQ SEQUENCE 143 AA; 15771 MW; AD4FBEFB10F6075C CRC64;
```

```
Query Match 16.0%; Score 325; DB 2; Length 143;
Best Local Similarity 37.5%; Pred. No. 8.7e-21;
Matches 90; Conservative 9; Mismatches 39; Indels 102; Gaps 5;
QY 3 FSLLLPRLECNCAISAHNRRLPGSSDSPASASPVAGITGMCTHARLLILYFFLVEMEFLLHV 62
Db 6 FLULSPKLECSGAILAHFNLCLLGSSDSPASASRVAGITGMHHRVRLI-FVFLVDTVY-Y 63
QY 63 VQAGLELPTSDPSVSASQSAARYTGHHARLCLANFCGRNRVSLMCPSSWSPELKQSTCL 122
Db 64 VQAGLELPTSD----- 76
QY 123 SLPKCKDYRRRAAVPGLFILFLRHRCPTLTQDEVQWCDHSSLPQSTPEIKHPPASASQVA 182
Db 77 -----LPASASQSA 85
QY 183 GTKDMHHVTWLIFFIENFLRQSLNSVTQAGVQWRNLGSLQPLPPGPKLFSCPSLLSSWD 242
Db 86 GITGVSHCTWAPFFFF-FLRQSF-IVIEAGVQWRDLSSLQLPPPGPKFSCGLGLPSSWD 143
Search completed: September 13, 2005, 19:34:51
Job time : 67 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 18:32:09 ; Search time 272 Seconds  
(without alignments)  
8674.673 Million cell updates/sec

Title: US-09-380-203-1  
Perfect score: 1442  
Sequence: 1 TTTTTCGATGGAG.....TTAAACAAAGCTTTAGACGA 1442

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2.6/prodata/1/ina/5A COMB.seq:\*
- 2: /cgn2.6/prodata/1/ina/5B COMB.seq:\*
- 3: /cgn2.6/prodata/1/ina/6A COMB.seq:\*
- 4: /cgn2.6/prodata/1/ina/6B COMB.seq:\*
- 5: /cgn2.6/prodata/1/ina/6C COMB.seq:\*
- 6: /cgn2.6/prodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1442	100.0	1442	2	US-08-454-557C-120
2	1442	100.0	1442	2	US-08-340-426D-120
3	1442	100.0	1442	2	US-08-450-673C-120
4	1442	100.0	1442	4	US-09-872-968-1
5	1223.4	84.8	1418	5	PCT-US95-17111A-120
6	1080.2	74.9	1381	2	US-08-454-557C-49
7	1080.2	74.9	1381	2	US-08-340-426D-49
8	1080.2	74.9	1381	2	US-08-450-673C-49
9	1080.2	74.9	1381	5	PCT-US95-17111A-49
10	590.4	40.9	42693	4	US-09-949-016-17317
11	590.4	40.9	42693	4	US-09-949-016-17318
12	597	36.5	107980	4	US-09-949-016-14370
13	519.6	36.0	39552	4	US-09-949-016-11799
14	506.4	35.1	13335	4	US-09-949-016-13162
15	502.4	34.8	43562	4	US-09-949-016-16222
16	499.2	34.6	100463	4	US-09-949-016-12511
17	499.2	34.6	100468	4	US-09-949-016-13725
18	493.2	34.2	84875	4	US-09-949-016-17334
19	493.2	34.2	84875	4	US-09-949-016-17335
20	493.2	34.2	84875	4	US-09-949-016-17336
21	493.2	34.2	84875	4	US-09-949-016-17337
22	493.2	34.2	85152	4	US-09-949-016-12665
23	493.2	34.2	85152	4	US-09-949-016-12666
24	493.2	34.2	85152	4	US-09-949-016-12667
25	493.2	34.2	85152	4	US-09-949-016-12668
26	492.6	34.2	17198	4	US-09-949-016-13776
27	482	33.4	44477	4	US-09-949-016-16767

C	28	480.6	33.3	37269	4	US-09-949-016-16672	Sequence 16672, A
	29	479.6	33.3	57914	4	US-09-949-016-11935	Sequence 11935, A
	30	479.6	33.3	57936	4	US-09-949-016-16921	Sequence 16921, A
	31	479.2	33.2	51719	4	US-09-918-686-2	Sequence 2, Appli
	32	479.2	33.2	92139	4	US-09-918-686-1	Sequence 1, Appli
	33	478.8	33.2	12091	4	US-09-949-016-15805	Sequence 15805, A
C	34	477.6	33.1	14452	4	US-09-949-016-13716	Sequence 13716, A
	35	477.6	33.1	60489	4	US-09-949-016-16287	Sequence 16287, A
	36	475.8	33.1	107980	4	US-09-949-016-14370	Sequence 14370, A
	37	475.4	33.0	14796	3	US-08-975-080-35	Sequence 35, Appl
	38	475.4	33.0	14796	3	US-09-630-706-10	Sequence 10, Appl
	39	475.4	33.0	14796	3	US-09-496-694B-3	Sequence 3, Appli
	40	475.4	33.0	14796	4	US-10-138-618-35	Sequence 35, Appl
	41	475.4	33.0	14796	4	US-09-918-186A-3	Sequence 3, Appli
	42	474.8	32.9	30867	4	US-09-949-016-12231	Sequence 12231, A
	43	474.8	32.9	30868	4	US-09-949-016-13452	Sequence 13452, A
C	44	473.2	32.8	28806	4	US-09-949-016-13217	Sequence 13217, A
	45	473	32.8	91279	4	US-09-949-016-15146	Sequence 15146, A

## ALIGNMENTS

RESULT 1  
US-08-454-557C-120  
; Sequence 120, Application US/08454557C  
; Patent No. 5830670  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; APPLICANT: Wands, Jack R.  
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection  
; TITLE OF INVENTION: of Alzheimer's Disease  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/454,557C  
; FILING DATE: 30-MAY-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ludwig, Steven R.  
; REGISTRATION NUMBER: 36,203  
; REFERENCE/DOCKET NUMBER: 0609.3840003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2540  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 120:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1442 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: both  
; MOLECULE TYPE: CDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 15...1139  
US-08-454-557C-120

Query Match 100.0%; Score 1442; DB 2; Length 1442;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTTCGATGGAGTTTCGCTCTTTGCCAGGCTGGAGTGCATGGCGCAA 60



Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	TTTTTTTTTTTTCAGATGGAGTTTTCGCTCTGCTGTCGCCAGGCTGAGTGCATGCGCAA	60
Db	1	TTTTTTTTTTTTCAGATGGAGTTTTCGCTCTGCTGTCGCCAGGCTGAGTGCATGCGCAA	60
QY	61	TCTCAGCTCACCGCAACCTCCGCTCTCCGCTTCAAGCGATTCTCTGCTCAGCTCC	120
Db	61	TCTCAGCTCACCGCAACCTCCGCTCTCCGCTTCAAGCGATTCTCTGCTCAGCTCC	120
QY	121	CAGTAGCTGGGATTAAGCATGTGACCCAGCTCGGCTAATTTGTAATTTTTTTAG	180
Db	121	CAGTAGCTGGGATTAAGCATGTGACCCAGCTCGGCTAATTTGTAATTTTTTTAG	180
QY	181	TAGAGATGAGTTTCTCCATGTTGTTCAGGCTGCTCGAATCTCCGACCTCAGATGATC	240
Db	181	TAGAGATGAGTTTCTCCATGTTGTTCAGGCTGCTCGAATCTCCGACCTCAGATGATC	240
QY	241	CCTCGTCTCGGCTCTCCCAAGTGTAGATACAGGACTGGCCACCATGCGGCTCTGCC	300
Db	241	CCTCGTCTCGGCTCTCCCAAGTGTAGATACAGGACTGGCCACCATGCGGCTCTGCC	300
QY	301	TGGCTAATTTTGTGGTAGAAACAGGTTTCACTGATGTGCCAAGCTGGTCTCTGAGC	360
Db	301	TGGCTAATTTTGTGGTAGAAACAGGTTTCACTGATGTGCCAAGCTGGTCTCTGAGC	360
QY	361	TCAAGCAGTCCACCTGCTCCAGCTCCCAAGTGTGAGTACAGGCGTGCAGCGTGC	420
Db	361	TCAAGCAGTCCACCTGCTCCAGCTCCCAAGTGTGAGTACAGGCGTGCAGCGTGC	420
QY	421	CTGGCTTTTTTATTTTTTATTTTTTAAAGACAGAGTGTCCCACTCTTACCCAGAGTGAAG	480
Db	421	CTGGCTTTTTTATTTTTTATTTTTTAAAGACAGAGTGTCCCACTCTTACCCAGAGTGAAG	480
QY	481	TGCAGTGTGTGATCAGAGTCACTGCGAGCTTCAACTCTGAGATCAAGATCTCTCTG	540
Db	481	TGCAGTGTGTGATCAGAGTCACTGCGAGCTTCAACTCTGAGATCAAGATCTCTCTG	540
QY	541	CCTCAGCTCCCAAGTGTGGGACCAAGATGACACCTACACCTGAGTGGCTAATTTTA	600
Db	541	CCTCAGCTCCCAAGTGTGGGACCAAGATGACACCTACACCTGAGTGGCTAATTTTA	600
QY	601	TTTTTATTTTTTAAATTTTTTGAACAGAGTCTCAACTCTGTACCCAGGCTGGAGTCA	660
Db	601	TTTTTATTTTTTAAATTTTTTGAACAGAGTCTCAACTCTGTACCCAGGCTGGAGTCA	660
QY	661	GGCGCAATCTGGCTCACTGCAACCTCTGCTCCCGGTTCAAGTTATCTCTCTGCCCA	720
Db	661	GGCGCAATCTGGCTCACTGCAACCTCTGCTCCCGGTTCAAGTTATCTCTCTGCCCA	720
QY	721	GCCTCTCAGTAGCTGGGACTACAGCGCCCAACAGCTGATCTCTGAGCTTGTATTTT	780
Db	721	GCCTCTCAGTAGCTGGGACTACAGCGCCCAACAGCTGATCTCTGAGCTTGTATTTT	780
QY	781	TAGTAGAGATGGGTTTACCACTGTTGCGCAGTGTGATCTGATCTCTGAGCTTGTATC	840
Db	781	TAGTAGAGATGGGTTTACCACTGTTGCGCAGTGTGATCTGATCTCTGAGCTTGTATC	840
QY	841	TGCTGCTCTGGCTCTCCCAAGTGTGCGATTTACAGGCTGAGCCACCAAGCGGCTTA	900
Db	841	TGCTGCTCTGGCTCTCCCAAGTGTGCGATTTACAGGCTGAGCCACCAAGCGGCTTA	900
QY	901	TTTTTAAATTTTGTGTTTGAATGGAATCTCACTGTGTTACCCAGGCTGGAGTGAAT	960
Db	901	TTTTTAAATTTTGTGTTTGAATGGAATCTCACTGTGTTACCCAGGCTGGAGTGAAT	960
QY	961	GGCCAAATCTCGGCTCACTGCAACCTCTGCGCTCCCGGCTCAAGGATTTCTCTGTCTCA	1020
Db	961	GGCCAAATCTCGGCTCACTGCAACCTCTGCGCTCCCGGCTCAAGGATTTCTCTGTCTCA	1020
QY	1021	GCCTCCCAAGCAGTGGGATTAAGGCACTTGGCCACCAACCCCGCTAATTTTGTATTT	1080
Db	1021	GCCTCCCAAGCAGTGGGATTAAGGCACTTGGCCACCAACCCCGCTAATTTTGTATTT	1080

RESULT 3

US-08-450-673C-120

; Sequence 120, Application US/08450673C

; Patent No. 5948888

; GENERAL INFORMATION:

; APPLICANT: de la Monte, Suzanne

; APPLICANT: Wanda, Jack R.

; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection

; TITLE OF INVENTION: of Alzheimer's Disease

; NUMBER OF SEQUENCES: 121

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

; STREET: 1100 New York Avenue, Suite 600

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20005-3934

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/450,673C

; FILING DATE: 30-MAY-1995

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Ludwig, Steven R.

; REGISTRATION NUMBER: 36,203

; REFERENCE/DOCKET NUMBER: 0609.3840004

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 371-2600

; TELEFAX: (202) 371-2540

; INFORMATION FOR SEQ ID NO: 120:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1442 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: both

; MOLECULE TYPE: cdna

; FEATURE:

; NAME/KEY: CDS

i	LOCATION:	15...1139
US-08-450-673C-120		
Query Match	100.0%; Score 1442;	DB 2; Length 1442;
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1442;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	TTTTTTTTTTTGAGATGGAGTTCGTCTTGTTGCCAGGCTGGAGTGCMAATGGCGCAA 60
Db	1	TTTTTTTTTTTGAGATGGAGTTCGTCTTGTTGGCCAGGCTGGAGTGCMAATGGCGCAA 60
Qy	61	TCTCAGCTCACCGCAACTCTCGCTCCCGGTTCAAGCATTCCTCGCTCAGCCTCCC 120
Db	61	TCTCAGCTCACCGCAACTCTCGCTCCCGGTTCAAGGATTCCTCGCTCAGCCTCCC 120
Qy	121	CAGTAGCTGGGATTACAGGCATGFCACCACAAGCTCGGCTAAATTTTGTATTTTTTTAG 180
Db	121	CAGTAGCTGGGATTACAGGCATGFCACCACAAGCTCGGCTAAATTTTGTATTTTTTTAG 180
Qy	181	TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGTCTCGAATCCCCAGCTCAGATGATC 240
Db	181	TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGTCTCGAATCCCCAGCTCAGATGATC 240
Qy	241	CTCTCGCTCGGCTCCCAAAGTCTAGATACAGGACTGGCCACCATGCCGGCTCTGCC 300
Db	241	CTCTCGCTCGGCTCCCAAAGTCTAGATACAGGACTGGCCACCATGCCGGCTCTGCC 300
Qy	301	TGGCTAATTTTTGTGGTAGAAA CAGGGTTTCACTGATGTGCCAAGCTGTCTCTCTGAGC 360
Db	301	TGGCTAATTTTTGTGGTAGAACAGGGTTTCACTGATGTGCCAAGCTGTCTCTCTGAGC 360
Qy	361	TCAAGCAGTCCACTGCTCAGCCTCCCAAAGTCTGGGATTACAGGCTGTCAGCCGTGC 420
Db	361	TCAAGCAGTCCACTGCTCAGCCTCCCAAAGTCTGGGATTACAGGCTGTCAGCCGTGC 420
Qy	421	CTGGCTTTTATATTTTATTTTTTTTAAGACACAGGTGTCCACCTCTTACCCAGGATGAAG 480
Db	421	CTGGCTTTTATATTTTATTTTTTTTAAGACACAGGTGTCCACCTCTTACCCAGGATGAAG 480
Qy	481	TGCAGTGTGTGATCACAGCTCAGTCAGAGCTTCAA CTCTCTGAGATCAAGCATCTCTCTG 540
Db	481	TGCAGTGTGTGATCACAGCTCAGTCAGAGCTTCAA CTCTCTGAGATCAAGCATCTCTCTG 540
Qy	541	CCTCAGCCTCCCAAGTAGCTGGGACCAAGA CATGCA CCACTACACCTGGCTAAATTTTAA 600
Db	541	CCTCAGCCTCCCAAGTAGCTGGGACCAAGA CATGCA CCACTACACCTGGCTAAATTTTAA 600
Qy	601	TTTTTATTTTTATTTTTTGACACAGAGTCTCAACTCTGTTCACCCAGGCTGGAGTGCAGT 660
Db	601	TTTTTATTTTTATTTTTTGACACAGAGTCTCAACTCTGTTCACCCAGGCTGGAGTGCAGT 660
Qy	661	GGGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTATTTCTCTCGCCCCA 720
Db	661	GGGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTATTTCTCTCGCCCCA 720
Qy	721	GCTCTCTGAGTAGCTGGGA CTA CAGGGGCCCA CACAGCCTAGCTAAATTTTTTTGTATTTTT 780
Db	721	GCTCTCTGAGTAGCTGGGA CTA CAGGGGCCCA CACAGCCTAGCTAAATTTTTTTGTATTTTT 780
Qy	781	TAGTAGAGATGGGGTTCCACATGTTTCG CAGGTTGATCTTGATCTCTGGACCTTGTGATC 840
Db	781	TAGTAGAGATGGGGTTCCACATGTTTCG CAGGTTGATCTTGATCTCTGGACCTTGTGATC 840
Qy	841	TGCTGCTCGGCTCCCAAAGTCTGGGATTACAGGCGTGAGCCA CACGCCCGGCTTA 900
Db	841	TGCTGCTCGGCTCCCAAAGTCTGGGATTACAGGCGTGAGCCA CACGCCCGGCTTA 900
Qy	901	TTTTTATTTTTTGTCTTTTGAATGGAATCTCACTCTGTGTACCAAGCTGGAGTGCNAAT 960
Db	901	TTTTTATTTTTTGTCTTTTGAATGGAATCTCACTCTGTGTACCAAGCTGGAGTGCNAAT 960
Qy	961	GGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGGCTCCAAAGGATTTCTCTGTGTCA 1020

Qy	241	CCTCGGCTCTCGGCTCCCAAAGTGCTAGATACAGGACTGGGCCACATGGCCCGGCTCTGCC	300
Db	241	CCTCGGCTCTCGGCTCCCAAAGTGCTAGATACAGGACTGGGCCACATGGCCCGGCTCTGCC	300
Qy	301	TGGCTAAATTTTGTGTGATAGAAACAGGGTTTCACTGATGTCGCCAAGCTGTGTTCTCTGAGC	360
Db	301	TGGCTAAATTTTGTGTGATAGAAACAGGGTTTCACTGATGTCGCCAAGCTGTGTTCTCTGAGC	360
Qy	361	TCAAGCAGTCCACTGTGCTCAGCCTCCCAAAGTGTGGGATTAACAGGCGTGCAGCGGTGC	420
Db	361	TCAAGCAGTCCACTGTGCTCAGCCTCCCAAAGTGTGGGATTAACAGGCGTGCAGCGGTGC	420
Qy	421	CTGGCCTTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	480
Db	421	CTGGCCTTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	480
Qy	481	TGCAGTGTGTGATCACAGCTCACTGCAGCCTTCAAATCTCTCGAGATCAAGCATCTCTCTG	540
Db	481	TGCAGTGTGTGATCACAGCTCACTGCAGCCTTCAAATCTCTCGAGATCAAGCATCTCTCTG	540
Qy	541	CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGTGGCTAAATTTT	600
Db	541	CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGTGGCTAAATTTT	600
Qy	601	TTTTTATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	660
Db	601	TTTTTATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	660
Qy	661	GGCGCAATCTTGGGCTCACTGCACCACTCTGCCTCCCGGGTTCAGATTATTTCTCTGCCCCA	720
Db	661	GGCGCAATCTTGGGCTCACTGCACCACTCTGCCTCCCGGGTTCAGATTATTTCTCTGCCCCA	720
Qy	721	GCCTCTCGAGTAGCTGGGACTACAGGCGGCCCAACAGCTGTGATCTCTGAGCTTGTGATCT	780
Db	721	GCCTCTCGAGTAGCTGGGACTACAGGCGGCCCAACAGCTGTGATCTCTGAGCTTGTGATCT	780
Qy	781	TAGTAGAGATGGGGTTCAACATGTTGCACAGTTGATCTCTGAGCTTGTGATCTCTGATC	840
Db	781	TAGTAGAGATGGGGTTCAACATGTTGCACAGTTGATCTCTGAGCTTGTGATCTCTGATC	840
Qy	841	TGCCTGCTCGGCTCCCAAAGTGTCTGGGATTAACAGGCGTGGAGCCACACGCCCGGCTTA	900
Db	841	TGCCTGCTCGGCTCCCAAAGTGTCTGGGATTAACAGGCGTGGAGCCACACGCCCGGCTTA	900
Qy	901	TTTTTAAATTTTGTGTTGAAATGGAATCTCACTCTGTGTACCGAGCTGGAGTGAAT	960
Db	901	TTTTTAAATTTTGTGTTGAAATGGAATCTCACTCTGTGTACCGAGCTGGAGTGAAT	960
Qy	961	GGCCAAATCTCGGCTCACTGCACCACTCTGCCTCCCGGGCTCAAGGATTTCTCTGTCTCA	1020
Db	961	GGCCAAATCTCGGCTCACTGCACCACTCTGCCTCCCGGGCTCAAGGATTTCTCTGTCTCA	1020
Qy	1021	GCCTCCCAAGCAGCTGGGATTAACGGGCACTTCGCCACCAACACCCCGCTAAATTTTGTATTT	1080
Db	1021	GCCTCCCAAGCAGCTGGGATTAACGGGCACTTCGCCACCAACACCCCGCTAAATTTTGTATTT	1080
Qy	1081	TCATTAGAGCGGGGTTTACCAATATTTGTGAGGCTGGTCTCAAACTCTCTGAGCTCAGGT	1140
Db	1081	TCATTAGAGCGGGGTTTACCAATATTTGTGAGGCTGGTCTCAAACTCTCTGAGCTCAGGT	1140
Qy	1141	GACCCACTGTGCTCAGCCTTCCAAAGTGTGGGATTAACGGGCTGGAGCCACTCACCCAG	1200
Db	1141	GACCCACTGTGCTCAGCCTTCCAAAGTGTGGGATTAACGGGCTGGAGCCACTCACCCAG	1200
Qy	1201	CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGCTCTTCTGATGTTGCCAGGCT	1260
Db	1201	CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGCTCTTCTGATGTTGCCAGGCT	1260
Qy	1261	GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGTAGGACCAACACCCAGCAGTCA	1320
Db	1261	GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGTAGGACCAACACCCAGCAGTCA	1320

Qy	1321	CAITTTTTAAACAGTTCACATCTTTATTTTATAGTATACATAGAACGTAATACAAATAAACAATGT	1380
Db	1321	CAITTTTTAAACAGTTCACATCTTTATTTTATAGTATACATAGAACGTAATACAAATAAACAATGT	1380
Qy	1381	CAAACTCGCAAAATTCAGTAGTAGTAACAGAGTTCCTTTTATAACTTTTAAACAAGCTTTTAGAG	1440
Db	1381	CAAACTCGCAAAATTCAGTAGTAGTAACAGAGTTCCTTTTATAACTTTTAAACAAGCTTTTAGAG	1440
Qy	1441	CA 1442	
Db	1441	CA 1442	

## RESULT 5

```

PCT-US95-17111A-120
; Sequence 120, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/POCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2500
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 120:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1418 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14..1418
PCT-US95-17111A-120

```



121 AGTAGGCTGGGATTAACAGGCTATGCA-CCAGCTCGGCTAAATTTTGTATTTTATTTAG 179  
 181 TAGAGATGAGATTTCTCCATGTTGGTCAAGCTGGTCTCGAATCTCCGACCTCAGATGATC 240  
 180 TAGAGATGAGATTTCTCCATGTTGGTCAAGCTGGTCTCGAATCTCCGACCTCAGATGATC 238  
 241 CTTCCGCTCTCGGCTCCCAAGTGTAGATACAGGACTGGCCACCAATGCCCGG-CTCTGC 299  
 239 CTTCCGCTCTCGGCTCCCAAGTGTAGATACAGGACTGGCCACCAATGCCCGGCTCTGC 298  
 300 CTGGCTAAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAG 359  
 299 CTGGCTAAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAG 358  
 360 CTCACGAGTCCAGCTGCTCAGCTCCCAAGTGTGGGATTAACAGGCTGGAGCCGTG 419  
 359 CTCACGAGTCCAGCTGCTCAGCTCCCAAGTGTGGGATTAACAGGCTGGAGCCGTG 418  
 420 CTGGGCTTTTATTTATTTTATTTTAAAGACACAGGTGCCCACTCTTACCCAGGATGAA 479  
 419 CTGGGCTTTTATTTATTTTATTTTAAAGACACAGGTGCCCACTCTTACCCAGGATGAA 478  
 480 GTGAGTGTGTGATCACAGCTCACTGAGCCCTTCAACTCTGAGATCAAGCATCTCCT 539  
 479 GTGAGTGTGTGATCACAGCTCACTGAGCCCTTCAACT-CTGAGATCAAGCATCTCCT 537  
 540 GCTCAGCTCCC-AGTAGCTGGGACCAAGACATGACCACTACACCTGGGTAAATTTT 598  
 538 GCTCAGCTCCCAGAGTAGCTGGGACCAAGACATGACCACTACACCTGGGTAAATTTT 597  
 599 TATTTTATTTTAAATTTTGTAGACAGAGTCTCAACTCTGTCACCCAGGCTGGATGCA 658  
 598 TATTTTATTTTAAATTTTGTAGACAGAGTCTCAACTCTGTCACCCAGGCTGGATGCA 657  
 659 GTGGGCAATCTTGGCTCACTGCAACTCTGCTCCCGGGTTCAGTTATTTCTCTGCC 718  
 658 GTGGGCAATCTTGGCTCACTGCAACTCTGCTCCCGGGTTCAGTTATTTCTCTGCC 717  
 719 CAGCTCTGAGTGTGGGACTACAGGCGCCCAACCGCTAGCTAAATTTTGTATTT 778  
 718 CAGCTCTGAGTGTGGGACTACAGGCGCCCAACCGCTAGCTAAATTTTGTATTT 777  
 779 TTTAGTAGAGATGGGG-TTCACCATCTGTCAGGCTTGAT-CTTGATCTGACCTCT 836  
 778 TTTAGTAGAGATGGGGTTTCAACATGTTCCAGGTTGATGATGATCTCTGACCTCT 837  
 837 GATCTGCTGCTCGGCTCCCAAGTGTGGGATTAACAGG-CTGAGGACCAACAGCCCG 895  
 838 GATCTGCTGCTCGGCTCCCAAGTGTGGGATTAACAGGAGCTGACGCCACCGCCG 897  
 896 GCTTATTTTAAATTTTGTGTTGTTGAAATGGAATCTCACTGTTTACCGAGCTGGAT 955  
 898 GCTTATTTTAAATTTTGTGTTGTTGAAATGGAATCTCACTGTTTACCGAGCTGGAT 957  
 956 GCAATGCGCAAAATCTCGGCTCACTGCAACTCTGCTCCCGGCTCAAGGATTTCTCC 1015  
 958 GCAATGCGCAAAATCTCGGCTCACTGCAACTCTGCTCCCGGCTCAAGGATTTCTCC 1017  
 1016 TCTCAGCTCCCAAGCAGCTGGGATTAACGGGACCTGCGCACCAACCGGCTAAATTTT 1075  
 1018 TCTCAGCTCCCAAGCAGCTGGGATTAACGGGACCTG-CAACACACCGGCTAAATTTT 1076  
 1076 TATTTTCAATAGAGCGGGGTTTCAACATATTTGTAGGCTGGTCTCAAACTCTGACCT 1135  
 1077 TATTTTCAATAGAGCGGGGTTTCAACATATTTGTAGGCTGGTCTCAAACTCTGACCT 1136  
 1136 CAGGTACCACTGCTCAGCTTCCAAAGTCTGCGATTAACAGCGGTGAGCCACTCA 1195  
 1137 CAGGTACCACTGCTCAGCTTCCAAAGTCTGCGATTAACAGCGGTGAG-CCGCTCA 1194  
 1196 CCAGCGGGCTAATTTAGATAAAAAAATATGTAGCAATGGGGGCTTCTGCTATGTTGCC 1255  
 1195 CCAGCGGGCTAATTTAGATAAAAAAATATGTAGCAATGGGGGCTTCTGCTATGTTGCC 1254

QY 1256 AGGCTGGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACCAACCCAGCC 1315  
 DB 1255 AGGCTGGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACCAACCCAGCC 1314  
 QY 1316 AGTCACATTTTAAACAGTTACATCTTTATTTAGTATCTAGAAAAGTAATACATAAA 1375  
 DB 1315 AGTCACA-TTTTAAACAGTTACATCTTTATTTAGTATCTAGAAAAGTATACGATAAC 1373  
 QY 1376 CATGTCAAACTCGCAATTCAGTAGTAAACAGATGTTCTTT 1414  
 DB 1374 ATGGCGGAACCTGCAAAATCGAGTAGTACAGAGTCTTTT 1412

RESULT 6  
 US-08-454-557C-49  
 ; Sequence 49, Application US/08454557C  
 ; Patent No. 5830670  
 ; GENERAL INFORMATION:  
 ; APPLICANT: de la Monte, Suzanne  
 ; APPLICANT: Wands, Jack R.  
 ; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection  
 ; NUMBER OF SEQUENCES: 121  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
 ; STREET: 1100 New York Avenue, Suite 600  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20005-3934  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/454,557C  
 ; FILING DATE: 30-MAY-1995  
 ; CLASSIFICATION: 514  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Ludwig, Steven R.  
 ; REGISTRATION NUMBER: 36,203  
 ; REFERENCE/DOCKET NUMBER: 0609.3840003  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 371-2600  
 ; TELEFAX: (202) 371-2540  
 ; INFORMATION FOR SEQ ID NO: 49:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1381 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: both  
 ; TOPOLOGY: both  
 ; US-08-454-557C-49

Query Match 74.9%; Score 1080.2; DB 2; Length 1381;  
 Best Local Similarity 94.4%; Pred. No. 2.9e-296;  
 Matches 1317; Conservative 0; Mismatches 53; Indels 25; Gaps 18;

QY 2 TTTTATTTTGTAGATGGAGTTTTCGCTCTTGTGTCAGGCTGGAGTGAATGGGGCAAT 61  
 DB 1 TTTTATTTTGTAGATGGAGTTTTCGCTCTTGTGTCAGGCTGGAGTGAATGGGGCAAT 60  
 QY 62 CTCAGCTCAACCGAACTCCGCTCCCGGGTTCAAGCGATTCTCTGCTCAGCTCCCC 121  
 DB 61 CTCAGCTCAACCGAACTCCGCTCCCGGGTTCAAGCGATTCTCTGCTCAGCTCCCC 120  
 QY 122 AGTAGCTGGATTAACGAGTGTGACCCAGCTGGCTAAATTTTGTATTTTATTTTACT 181  
 DB 121 AGTAGCTGGATTAACGAGTGTGCA-CCACCTCGGCTAAATTTTGTATTTTATTTAGT 179  
 QY 182 AGAGATGAGATTT--CTCCATCTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGAT 239











```

QY 645 CAGGCTGGAGTCAGTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGTTCAAG 704
Db 41487 CAGGCTGGAGTCAGTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGTTCAAG 41428
QY 705 TTAATCTCTGCTCCCGCAAGCTCTGAGTGGAGTACAGGGCCCAACAGCCCTAGCT 764
Db 41427 CAATCTCTGCTCCCGCAAGCTCTGAGTGGAGTACAGGGCCCAACAGCCCTAGCT 41368
QY 765 AATTTTTTTTGTATTTTGTAGATGGGG-TTCACCATGTTGCGCAGGTTCATCTTGAT 823
Db 41367 GCTAAATTTTGTATTTTGTAGATGGGG-TTCACCATGTTGCGCAGGTTCATCTTGAT 41308
QY 824 CTCTGAGCT--TGTGATCTGCTCTGCTCCCGCTCCCAAGTGTGGGATTACAGCGTG 881
Db 41307 CTCTGAGCTCAGGAGATCTCTGCTCCCGCTCCCAAGTGTGGGATTACAGCGTG 41248
QY 882 AGCCACACCGCCGGCT--ATTTTTTAATTTTGTGTTTGTGAATGAATCTCACTCT 938
Db 41247 AGCCACACCGCCGGCT--ATTTTTTAATTTTGTGTTTGTGAATGAATCTCACTCT 41188
QY 939 GTTACCCAGGCTGGAGTCAATGGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGG 998
Db 41187 GTTACCCAGGCTGGAGTCAATGGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGG 41128
QY 999 CTCACGAGATCTCTGCTCTGAGCTCCCAAGCAGCTGGGATTACAGGGCACTGCCACCA 1058
Db 41127 CTCACGAGATCTCTGCTCTGAGCTCCCAAGCAGCTGGGATTACAGGGCACTGCCACCA 41068
QY 1059 CACCCGCTAAATTTTGTATTTTTCATAGAGGGGGTTTCAACATATTTGTCAGGCTGG 1118
Db 41067 CACCCGCTAAATTTTGTATTTTTCATAGAGGGGGTTTCAACATATTTGTCAGGCTGG 41008
QY 1119 TCTCAACTCTGAGCTCAGGTGACCCACCTGCTCAGCTTCCAAAGTCTGGGATTAC 1178
Db 41007 TCTCAACTCTGAGCTCAGGTGACCCACCTGCTCAGCTTCCAAAGTCTGGGATTAC 40948
QY 1179 AGCGTGAGCCACCTCACCAGCC-----GGCTAAATTTAGATAAAAAAATATAGCAATG 1234
Db 40947 AGCGTGAGCCACCTCACCAGCC-----GGCTAAATTTAGATAAAAAAATATAGCAATG 40888
QY 1235 GGGGCTCTGATGTTGGCCAGGCTGGTCTCAACTCTGCTCCAAAGTCTGGGATTAC 1294
Db 40887 GGGGCTCTGATGTTGGCCAGGCTGGTCTCAACTCTGCTCCAAAGTCTGGGATTAC 40828
QY 1295 AATGAGCCAC 1304
Db 40827 TCTCAGCCTC 40818

```

```

RESULT 11
US-09-949-016-17318/c
; Sequence 17318, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17318
; LENGTH: 42693
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17318

```

```

Query Match 40.9%; Score 590.4; DB 4; Length 42693;
Best Local Similarity 70.8%; Pred No. 1.3e-156;
Matches 941; Conservative 0; Mismatches 346; Indels 43; Gaps 10;

QY 4 TTTTGTGAGATGAGTGGTTCGCTCTTTGTTGCCAGGCTGAGTGCATATGGCGCAATCT 63
Db 42133 TTTTGTGAGATGAGTGGTTCGCTCTTTGTTGCCAGGCTGAGTGCATATGGCGCAATCT 42075
QY 64 CAGCTCACCGCAACCTCCGCTCCCGGTTCAAGCATCTCTGCTCAGCTCAGCTCCCGAG 123
Db 42074 CAGCTCACCGCAACCTCCGCTCCCGGTTCAAGCATCTCTGCTCAGCTCAGCTCCCGAG 42015
QY 124 TAGCTGGGATTAACAGGCATGTGCACCCAGCTCGGCTAATTTTGTATTTTGTATAGTAG 183
Db 42014 TAGCTGGGATTAACAGGCATGTGCACCCAGCTCGGCTAATTTTGTATTTTGTATAGTAG 41959
QY 184 AGATGGAGTTTCTCCATGTTGGTCAAGCTGCTCGAACTCCCGACCTCAGATGATCCCT 243
Db 41958 AGATGGAGTTTCTCCATGTTGGTCAAGCTGCTCGAACTCCCGACCTCAGATGATCCCT 41899
QY 244 CCGTCTCGGCTCCCGCAAGTGTAGTACAGGACTGGCCACCATGCCCGCTCTGCTGG 303
Db 41898 CCGTCTCGGCTCCCGCAAGTGTAGTACAGGACTGGCCACCATGCCCGCTCTGCTGG 41840
QY 304 CTAAATTTTGTGTAGAAACAGGGTTTCTCATGATGTGCCAAGCTGCTCTCTGAGCTCA 363
Db 41839 AGTATGAAGCAGTACTCCCGGATATGCAAGGCTCTGAAACCTCTGAGACTC 41780
QY 364 AGCAGTCAACCTGCTCAGCTCCCGCAAGTGTGGGATTTACAGGCTGCGAGCTGCTG 423
Db 41779 TTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 41720
QY 424 GCC-----TTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 465
Db 41719 CCCCCCTCATCTCCCTGCTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 41660
QY 466 TTATCCAGAGTGAAGTGCAGTGGTGTGATCAGCTCAGCTCAGCTGAGCTCTCACTCTGAGA 525
Db 41659 TTATCCAGAGTGAAGTGCAGTGGTGTGATCAGCTCAGCTCAGCTGAGCTCTCACTCTGAGA 41600
QY 526 TC-AAGCATCTCTGCTCAGCTCCCGCAAGTGTGGGATTTACAGGCTGCGAGCTGCTG 584
Db 41599 TC-AAGCATCTCTGCTCAGCTCCCGCAAGTGTGGGATTTACAGGCTGCGAGCTGCTG 41540
QY 585 ACCTGGCTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 644
Db 41539 GCCCCAACTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 41488
QY 645 CAGGCTGGAGTGCAGTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGTTCAAG 704
Db 41487 CAGGCTGGAGTGCAGTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGTTCAAG 41428
QY 705 TTAATCTCTGCTCCCGCAAGCTCTGAGTGGAGTACAGGGCCCAACAGCCCTAGCT 764
Db 41427 CAATCTCTGCTCCCGCAAGCTCTGAGTGGAGTACAGGGCCCAACAGCCCTAGCT 41368
QY 765 AATTTTTTTTGTATTTTGTAGTAGATGGGG-TTCACCATGTTGCGCAGGTTCATCTTGAT 823
Db 41367 GCTAAATTTTGTATTTTGTAGTAGATGGGG-TTCACCATGTTGCGCAGGTTCATCTTGAT 41308
QY 824 CTCTGAGCT--TGTGATCTGCTCTGCTCCCGCTCCCAAGTGTGGGATTACAGCGTG 881
Db 41307 CTCTGAGCTCAGGAGATCTCTGCTCCCGCTCCCAAGTGTGGGATTACAGCGTG 41248
QY 882 AGCCACACCGCCGGCT--ATTTTTTAATTTTGTGTTTGTGAATGAATCTCACTCT 938
Db 41247 AGCCACACCGCCGGCT--ATTTTTTAATTTTGTGTTTGTGAATGAATCTCACTCT 41188
QY 939 GTTACCCAGGCTGGAGTCAATGGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGG 998
Db 41187 GTTACCCAGGCTGGAGTCAATGGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGG 41128

```

```
QY 999 CTCACGCGATTCTCTGCTCTCAGCTCCCAAGCGCTGGATTACGGGCACTGCCACA 1058
Db 41127 TTCAAGCGATTCTCTGCTCTCAGCTCCCAAGTAGTAGTACAGGCGATGTGCCACA 41068
QY 1059 CACCCGCTAAATTTTGTATTTTTCATTAGAGCGGGTTTCAACCATATTTGTGAGCTGG 1118
Db 41067 CGCCCGCTAAATTTTGTATTTTGTAGAGAGAGTATTTACCATTTGGCCAGACTGG 41008
QY 1119 TCTCAAACTCTGACCTCAGGTGACCCACTGCTCAGCTTCCAAAGTGTGGGATTAC 1178
Db 41007 TCTTGAATCTCTAACTTCAAGTGTGATCCACCGCTCTGCTCCCAAGTGTGGGATTAC 40948
QY 1179 AGCGGTGAGCCACTCACCCAGC---GGCTAAATTTAGATAAAATAATATGCAATG 1234
Db 40947 AGGCATGAGCCACCACCTGGCTCTGCCAACTAAATTAATAAAATTTTGTAGAGA 40888
QY 1235 GGGGGTCTGTATGTGCGCCAGGCTGGTCTCAAACTTCTGGCTCATGCAATCTTCCA 1294
Db 40887 CAGGATCTCACTATGTGTCCAGACTGGTCTTGAACCTCTGGCTTAAGGATCGTCTG 40828
QY 1295 AATGAGCCAC 1304
Db 40827 TCTCAGCCTC 40818
```

## RESULT 12

```
US-09-949-016-14370/c
; Sequence 14370, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: PASTSEQ for Windows Version 4.0
; SEQ ID NO 14370
; LENGTH: 107980
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(107980)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14370
```

```
Query Match 36.5%; Score 527; DB 4; Length 107980;
Best Local Similarity 71.2%; Pred. No. 2e-138;
Matches 915; Conservative 0; Mismatches 280; Indels 90; Gaps 13;
```

```
QY 1 TTTTCTTTTTCAGATGAGTCTTTCGCTCTGTGTGCGCCAGGCTGAGTGAATGGCGCAA 60
Db 22848 TATTTATTTTATGACACAGAGTCTTGCTCTGTCTGTCACCCAGGCTGAGTGGCACGG 22789
QY 61 TCTCAGCTCACCGCAACTCCGGCTCCGGGTTCAAGCGATTCTCTGCTCAGCCTCCC 120
Db 22788 TCTCGGTTCACTGCAACCTCCGGCTCCAGGTTCAAAATGATTCTCTGCTTTAGCCTCT 22729
QY 121 CAGTAGCTGGATTACAGCATGTGACCCAGCTCGGCTAAATTTTGTATTTTGTAG 180
Db 22728 GAGTAGCTGGATTACAGGACGACCAACACCGCAGGCTAAATTTT---TGTATTTTGTAG 22672
QY 181 TAGAGATGAGGTTTCTCCATTTGTTGGTTCAGGCTGGTCTCGAAGTCCCGACCTCAGATGATC 240
Db 22671 TAGAGACGGGGTTTACCAATTTTGGCCAGGCTGGTCTTGAATTTCCGACTTC---GTGATC 22614
```

```
QY 241 CCTCGCTCTCGGCTCTCCAAAGTGTAGATACAGGACTGGCCACCATGCCGGCTCTGCC 300
Db 22613 CGCTGCTCTCAGCTCTCCAAAGTGTCT-----GGGATTACAGGCTGAGCCACCACT 22560
QY 301 TGGCTAAATTTTGTGTAGAACAGGGTTTCACTGATGTGCGCAAGCTG----- 349
Db 22559 GGCCTATTTATTTTGTAGATGAGTCTTGTCTGTGTGCCAGGCTGGAGTGCAGTGG 22500
QY 350 -----GTCTCTGAGCTCAAGCAGTCAACCTGCTCAGCTCCAGC-----GTGCCCTGGCC 426
Db 22499 TGCCATCTTGGCTCACTCAACCTCTGGGCTCTGGGTTTCAAGTGATTCTCTGCTCCAGC 22440
QY 384 CTCCTCAAGTGTGGGATTACAGGCGTGCAGC-----GTGCCCTGGCC 426
Db 22439 GTCCCGAGTAGCTGGGATTACAGGCGCAAGCCACCATGCCGGCTAAATTTTGTATTTCT 22380
QY 427 TTTTATTTTATTTTAAATTTTAAAGACACAGTGTCCCACTCTTACCCAGGATGAAGTGCAGT 486
Db 22379 TTTTATTTTATTTTGTAGATGAGTCTCAGTCTCTTGTGCCAGGCTGGAGTGCAC 22320
QY 487 GGTGTGATCAGAGCTCACTGAGCCTTCAACTCTCTGAGATCAAGC-ATCCTCTCTCCCTCA 545
Db 22319 AGTGGATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGCGATTCTCTCGCTCA 22260
QY 546 GCCTCCCAAGTAGCTGGGACCAAGACATGACACCACTACACTGCTGCTTAATTTTAT----- 601
Db 22259 GCCTCCGAGTAGCTGGGATTACAGGCGCAAGCCACCATGCCGGCTAAATTTTGTATTT 22200
QY 602 -TTTTATTTTAAATTTTGTAGACAGAGTCTCAACTCTGTCAACCAGGCTGGAGTGCAGT 660
Db 22199 CTCTCTCTTTTATTTTGTAGATGAGTCTCAGTCTTGTCACTGAGTGGAGTGCAC 22140
QY 661 GCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTCCGCCCA 720
Db 22139 GGTGGATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGCGATTCTCTCGCTCA 22080
QY 721 GCCTCTGAGTAGCTGGGACTACAGGCGCCACCAAGCTAGCTAAT-TTTTGTATTT 779
Db 22079 GCCTCTTGAATACCTGGGATTATAGATGCTGCGCCACCATGCCGGCTAAATTTTGTATTT 22020
QY 780 TTAGTAGAGA-TGGGGTTTCACTGTCGCGAGGTGTGATCTTGTACTCTGAGCTCT--TGT 836
Db 22019 GTAGTAGAGACAGGTTTCAACATATTGAAAGGCTGTCTCGAATCTCTGACCTCAGT 21960
QY 837 GATCTGCTGCTCGGCTCCCAAGTGTGGGATTACAGGCTGAGCCACCAAGCCCGG 896
Db 21959 GATCTGCTGCTCAGCCTCCCAAGTGTGGGATTACAGGCTGAGCCACCATGCCAG 21900
QY 897 CTTAT-----TTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTATCCC 945
Db 21899 CTTTTTACTTTTATTTATTTTATTTATTTATTTGAAACAGAGTCTCACTCTGTACATC 21840
QY 946 AGGCTGGAGTGAATGGCCAAATCTCGCTCACTGCAACCTCTGCTCCCGGGCTCAAGC 1005
Db 21839 AGGCTGGAGTGAATGGCCAAATCTCGCTCACTGCAACCTCTGCTCCCGGGTCAAGC 21780
QY 1006 GATCTCTGCTCAGCTCCCAAGCAGTGGGATTACAGGCACTTGCACCACTGCGCCAG 1065
Db 21779 GATCTCTGCTCAGCCTCAGCCTGAGTAGCTGGGACTATAGGCACTGCGCCAGCCAG 21720
QY 1066 CTAATTTTGTATTTTATTTAGAGGGGGTTTACCATATTTGTGAGCTGGTCTCATA 1125
Db 21719 CTGATTTTGTG--TTTGTAGTAGAGAGGTTTACCATATTTGTGTCAGGCTGGTCTTGA 21662
QY 1126 CTCTGAGCTCAGGTGACCCACTCTGCTCAGCTTTCCAAAGTGTGGGATTACAGGCTG 1185
Db 21661 CCGCTGAGCTCA--TGATCGGCTGCTTGGCTCCCAAGTGTCTCGGATTACAGGCTG 21604
QY 1186 AGCCACCTCACCCAGCGGCTAATTT 1210
Db 21603 AGCCACCCAGCCCTAATTTT 21579
```



Query Match	36.0%	Score 519.6	DB 4	Length 39552
Best Local Similarity	69.0%	Pred. No. 1.5e-136		
Matches 871	Conservative 0	Mismatches 1349	Indels 42	Gaps 10
Qy 1	TTTTTTTTTTGAGATGGAGTTTTGCTCTGTGTGGCCACAGGCTGAGATGCAATGCGCAA	60		
Db 34421	TTTTTTTTTTTGGAGATGGAG-TTTCACTCTGTGTGGCCACAGGCTGAGCGCAATGGCGTGA	34363		
Qy 61	TCTCAGCTCACCGCAAACCTCGGCTCCGGGTTCGAAGCGAMTCTCTCGCTCAGCCTCCC	120		
Db 34362	TCTCCATTACCGCAAACCTCGGCTTCAGGTTCAACGATTTCTCTCGCTCAGCCTCCC	34303		
Qy 121	CAGTAGCTGGGATTCAGGCAATGTGCACCCAGCTCGGCTAATTTTGTGATTTTTTTTAG	180		
Db 34302	AAGTAGCTGGGACTCAGGCAATGTGCACCACTGGGCTAATTTTGTGATTTT	34247		
Qy 181	TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC	240		
Db 34246	TAGAGACGGGTTTCTCCATGTTGGTCAGGCTGGTCTTGAACTCCCAACCTCAGGTGATT	34187		
Qy 241	CCTCGCTCTCGGCTCCCAAAGTGCTAGATACAGGACTGGCCACCATGCCGGCTCTGCC	300		
Db 34186	CAGTGGCTCTCGGCTCTCTAAGT-AATTGAGATTTACAGGCGTGAGCACCGGCTGGGCT	34128		
Qy 301	TGGCTAATTTTGTGGTAGAAACAGGGTTTTCACTGATGTGCCCAAGCTGGTC-----TC	354		
Db 34127	AAATTTTGTATTTTAAATAGAGATGGGTTTTGGCCATGTTGGCCAGCTGGTCTCCAATTC	34068		
Qy 355	CTGAGCTCAAGCAGTCCACTGGCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCA	414		
Db 34067	CTGACCTCAAGTGATCCACCGGCTTGCTCTCCCAAAGTGTGGGATTTACAGGCGTGATC	34008		
Qy 415	CCGTGGCTGGCCTTTTATTTATTTTATTTTAAAGCACAGGTGTCCCACTCTTACCCAGG	474		
Db 34007	CACATGCCCAGGCTTCATTTCTTTTCAAAGCTGATATATATTTCACTATATGTATAACC	33948		
Qy 475	ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCGAGCCTTCAACTCTCGAGATCAAGCATC	534		
Db 33947	ACATTGTTATCTATCAATCTGATGATAGTACTCTGGGTGTTTCCAAATTTTAAATTTT	33888		
Qy 535	CTCCTGCTCAGGCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGGCTAA	594		
Db 33887	TAATTTATTTTATAGTTTCAATTTTATTTTAA-----TTTTCT	33852		

US-09-949-016-13162

```
Query Match      35.1%; Score 506.4; DB 4; Length 13335;
Best Local Similarity 71.0%; Pred. No. 5e-133;
Matches 887; Conservative 0; Mismatches 306; Indels 57; Gaps 14;

QY 1 TTTTATTTTTCAGATGGAGTTTCGCTCTCTGTCGCCAGAGCTGGAGTGCATGGCGAA 60
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 61 TCTCAGCTCAGCGCAACCTCCGCCCTCCCGGGTTCAAGCGATTCTCTGCTCAGGCTCCC 120
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 3426 TATTGCCCTCAGTGAACCTCCGCCCTCCCGGGTTCAAGCGATTCTCTGCTCAGGCTCCC 3485
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 121 CAGTAGCTGGGATTCAGGCAATGTCACCCAGCTCGGCTAAATTTGTAATTTTTTTTAG 180
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 3486 GAGTAGCTGGGATTCAGGCAATGTCACCCAGCTCGGCTAAATTTGTAATTTTTTTTAG 180
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 181 TAGAGATGGAGTTTCTCCATGTTGGTTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 3544 TAGAGACAAGGTTTTCACCGTGTAGCCAGGATGGTCTCTATCTCTGACCT--TGTGATC 3601
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 241 CCTCGCTCTCGGCCCTCCCAAAGTGTGTAGATACA---GGACTGGCCACCATGCCGGCTCT 297
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 3602 CGCCCGCTCAGGCTCCCAAGTGTGTAGATACA---GGACTGGCCACCATGCCGGCTCT 3661
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 298 GCCTGGCTAAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGGT----- 351
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 3662 -----TGTATTTTAGTAGAAGGGTTTTCACCATATTTGGCCAGGCTGGTTTCAA 3712
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 352 CTCTCAGCTCAGGCACTCCAGCTCCAGCTCCCAAGTGTGGGATTCACAGCGTG 411
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 3713 CTCTCGGCTCAAGTGTACCACTCGCTCTGCTGCCAAGTGTGGGATTAATAGTGTG 3772
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 412 CAGCGCTGCTCGGCTTTTATTTATTTTAAAGACACAGGTGTCCCACTCTTACCC 471
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 3773 AGCCACCATGTCTGGCCACTGTGATCTCAATTTTCAATTTTTTTTTTTTACTTTTGTG 3832
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 472 AGATGAAGTGTGAGTGTGATCTACAGCTCACTGAGGCTTCAACTCTGATGATCAAG- 530
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 3833 GAGATGGTCTACAGGCTGT-----CTTGAACCTCTCGGCTCAAGT 3873
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 531 CATCTCTCGCTCAGCTCCCAAGTGTGGGACCAAGACATGCACCCTACACCTGG 590
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 3874 GATCTCCCACTCAGCTCCCTCAAGTGTGGGATTCAGGATGAGCCACTACACCGAG 3933
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 591 CTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 650
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 3934 CCAA-CTTGTCTTTTTTTTTTTTTTTTTTTTTTTTGGAGACGGAGTCTC-GCTCTGTC-CCCAGGTT 3990
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 651 GGAGTGCAGTGGGCAATCTTGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTC 710
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 3991 GGAGTGCAGTGGGCAATCTTGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTC 4050
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 711 TCTCGCCAGCTCTCTGAGTAGCTGGGACTACAGCGCCACACCGCTAGC--TAATTT 768
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 4051 TCTCGCTCAGGCTCCCGAGTAGCTGGGACTACAGCGCCACACCGCTAGC--TAATTT 4110
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 769 TTTTGTATTTTATTTAGTAGATGGG-GTTTCAACATGTTTCGCCAGGTTGATCTTGTATCTCT 827
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 4111 TTTTGTATTTTATTTAGTAGACGGGATTTTTCACCGTGTAGCCAGGATGTCTCGATCTCC 4170
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 828 GGAACCTGTGATCTGCTCGCTCCCGCTCCCAAGTGTGGGATTCAGGCGGTGAGCCAC 887
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 4171 TGACCTCTGATCTCGCCCGCTCCCGCTCCCAAGTGTGGGATTCAGGCGGTGAGCCAC 4230
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 888 CAGCGCCGCG-----TTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 941
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 4231 TGGCCCGGCGCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATAGGAGTCTCAGTCTGTC 4290
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 942 ACCCAGGCTGGAGTGAATGGCAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTC 1001
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425

QY 4291 TCTCAGGCTGGAGTGGTGTGATCTCAGTTCACTGCAACCTCCCGCTCCTGAGTTT 4350
Db TTTTATTTTTCAGATGGAGTTCACACTCTGTGCCCCAGGCTGGAGTGCATGGCGAA 3425
```

## RESULT 15

```
US-09-949-016-16222/c
; Sequence 16222, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16222
; LENGTH: 43562
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(43562)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16222
```

```
Query Match      34.8%; Score 502.4; DB 4; Length 43562;
Best Local Similarity 69.4%; Pred. No. 1.2e-131;
Matches 861; Conservative 0; Mismatches 326; Indels 53; Gaps 11;

QY 1 TTTTATTTTTCAGATGGAGTTTCGCTCTTGTTCGCCAGGCTGAGTGCATGGCGAA 60
Db TTTTATTTTTCAGATGGAGTTT--GCTCTGTTCGCCAGGCTGAGTGCATGGCGAA 6716

QY 61 TCTCAGCTCAGCGCAACCTCCGCCCTCCCGGGTTCAAGCGATTCCTCGCTCAGGCTCCC 120
Db TTTTATTTTTCAGATGGAGTTT--GCTCTGTTCGCCAGGCTGAGTGCATGGCGAA 6716

QY 6715 TCTCGGCTCACTGCAACCTCCGCCCTCCCGGGTTCAAGCTATTCTCTCGCTCAGGCTCCT 6656
Db TTTTATTTTTCAGATGGAGTTT--GCTCTGTTCGCCAGGCTGAGTGCATGGCGAA 6716

QY 121 CAGTAGCTGGGATTCAGGCAATGTCACCCAGCTCGGCTAAATTTTGTATTTTTTTTAG 180
Db TTTTATTTTTCAGATGGAGTTT--GCTCTGTTCGCCAGGCTGAGTGCATGGCGAA 6716

QY 6655 GAGTAGCTGGGATTCAGGCAATGTCACCCAGCTCGGCTAAATTTTGTATTTTTTTAG 6600
Db TTTTATTTTTCAGATGGAGTTT--GCTCTGTTCGCCAGGCTGAGTGCATGGCGAA 6716

QY 181 TAGAGATGGAGTTTCTCCATGTTGGTTCAGGCTGGTCTCGAATCTCCGACCTCAGATGATC 240
Db TTTTATTTTTCAGATGGAGTTT--GCTCTGTTCGCCAGGCTGAGTGCATGGCGAA 6716

QY 6599 TAGAGACAGGGTTTTCACCGTGTGGTTCAGGCTGGTCTCTAACTCTGACCTCAGATGATC 6540
Db TTTTATTTTTCAGATGGAGTTT--GCTCTGTTCGCCAGGCTGAGTGCATGGCGAA 6716

QY 241 CCTCGCTCTCGGCTCCCAAGTGTAGATACAGGACTG---GCCACCATGCCCGCTCT 297
Db TTTTATTTTTCAGATGGAGTTT--GCTCTGTTCGCCAGGCTGAGTGCATGGCGAA 6716

QY 6539 TGACCGCTCTGCGCTCCCAAGAGTCTGGGATTCAGGCGGTGAGCCATCATGCCACGCCA 6480
Db TTTTATTTTTCAGATGGAGTTT--GCTCTGTTCGCCAGGCTGAGTGCATGGCGAA 6716

QY 298 GCTCGGCTAAATTTTGTGTGTAAGAACAGGTTTTCATGATGTGCCAACAGCTGGTCTCTCTG 357
Db TTTTATTTTTCAGATGGAGTTT--GCTCTGTTCGCCAGGCTGAGTGCATGGCGAA 6716
```



[illegible]

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 14:57:42 ; Search time 5116 Seconds  
(without alignments)  
10728.827 Million cell updates/sec

Title: US-09-380-203-1  
Perfect score: 1442  
Sequence: 1 TTTTITTTTTTGAGATGAG.....TTAAACAAAGCTTGAAGCA 1442

Scoring table: IDENTITY\_NUC  
Gapop\_10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsa1:\*  
9: gb\_gsa2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	521.8	36.2	622	CA437412	CA437412 UI-H-DT0-
2	460.6	31.9	3016	BC068461	BC068461 Homo sapi
3	446.8	31.0	549	BM996475	BM996475 UI-H-DT0-
4	396.2	27.5	1842	BC011119	BC011119 Homo sapi
5	391.4	27.1	877	AQ739838	AQ739838 HS_5505_A
6	384.6	26.7	4191	AL832992	AL832992 Homo sapi
7	383.6	26.6	1345	AV762220	AV762220 AV762220
8	379.9	26.3	910	BQ722917	BQ722917 AGENCOURT
9	378.4	26.2	767	AV700498	AV700498 AV700498
10	377.6	26.2	613	CD709038	CD709038 EST25565
11	377.6	26.2	2097	CR614786	CR614786 full-length
12	377.4	26.2	604	CA439577	CA439577 UI-H-DT0-
13	376.4	26.1	628	CD690152	CD690152 EST6675 h
14	375.8	26.1	820	BUS68843	BUS68843 AGENCOURT
15	375.4	26.0	2777	HS8802772	HS8802772 Homo sapi
16	374.8	26.0	782	EX644719	EX644719 DKF2p681N
17	374.2	26.0	600	CA439719	CA439719 UI-H-DT0-
18	373.8	25.9	658	AQ393450	AQ393450 CITR1-E1-
19	373.2	25.9	911	AQ746594	AQ746594 HS_2278_A
20	370.2	25.7	682	EX481615	EX481615 DKF2p686A
21	370.2	25.7	637	BQ438614	BQ438614 AGENCOURT
22	368.8	25.6	970	BM468547	BM468547 AGENCOURT
23	367.6	25.5	5556	AQ839814	AQ839814 260L13-C5
24	367.4	25.5	4087	BC024593	BC024593 Homo sapi

25	366.8	25.4	3143	3	HS8805999	BS537892 Homo sapi
26	366.6	25.4	839	1	AV700988	AV700988 AGENCOURT
27	366.2	25.4	895	5	BU603620	BU603620 AGENCOURT
28	365.8	25.4	660	5	EX642324	EX642324 DKF2p686M
29	365.2	25.3	891	5	BU174155	BU174155 AGENCOURT
30	364.8	25.3	660	5	EX484570	EX484570 DKF2p686I
31	364.2	25.3	1654	3	CR591254	CR591254 full-length
32	364.2	25.3	671	4	BM722145	BM722145 UI-B-ROO-
33	363.8	25.2	2044	3	BC037797	BC037797 Homo sapi
34	363.2	25.2	841	7	CF994276	CF994276 AGENCOURT
35	363	25.2	784	9	CL423322	CL423322 RP11-294J
36	361.4	25.1	611	5	BM996878	BM996878 UI-H-DHO-
37	360.8	25.0	820	5	BU852195	BU852195 AGENCOURT
38	360.8	25.0	930	5	BQ884038	BQ884038 AGENCOURT
39	360	25.0	1148	1	AV761207	AV761207 AV761207
40	359.8	25.0	667	9	AG150197	AG150197 Pan trogl
41	359.4	24.9	870	5	BQ711438	BQ711438 AGENCOURT
42	358.8	24.9	887	9	AG186062	AG186062 Pan trogl
43	357.8	24.8	679	2	BF346320	BF346320 602018525
44	357.6	24.8	888	5	BU594689	BU594689 AGENCOURT
45	357.4	24.8	722	8	AQ583981	AQ583981 RPC1-11-4

## ALIGNMENTS

RESULT 1  
CA437412  
LOCUS CA437412 622 bp mRNA linear EST 08-NOV-2002  
DEFINITION UI-H-DT0-avk-a-06-0-UI.s1 NCI CGAP DT0 Homo sapiens cDNA clone  
UI-H-DT0-avk-a-06-0-UI 3', mRNA sequence.  
ACCESSION CA437412  
VERSION CA437412.1 GI:24801832  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 622)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
TITLE Unpublished (1997)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Dr. Jose Mercuende  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, bento-soares@uiowa.edu  
The following repetitive elements were found in this cDNA  
sequence: 11-305, >ALU (matched complement) 221-440, >ALU (matched  
complement) 452-620, >ALU (matched complement)  
Seq primer: M13 FORWARD  
POLYA=Yes.

## FEATURES

Location/Qualifiers  
1..622  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="UI-H-DT0-avk-a-06-0-UI"  
/tissue\_type="Metastatic Chondrosarcoma"  
/dev\_stage="Adult"  
/clone\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI CGAP DT0"  
/notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
NCI CGAP DT0 is a cDNA library containing the following  
tissue(s): Metastatic Chondrosarcoma in Lung. The library  
was constructed according to Bonaldo, Lennon and Soares,  
Genome Research, 6:791-806, 1996. First strand cDNA  
synthesis was primed with an oligo-dT primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AACTGTTCCG.

TAG\_TISSUE=lung metastatic chondrosarcoma  
TAG\_LIB=UI-H-270  
TAG\_SEQ=AACTGTTCCG

ORIGIN

Query Match 36.2%; Score 521.8; DB 6; Length 622;  
Best Local Similarity 95.9%; Pred. No. 7.1e-67;  
Matches 590; Conservative 0; Mismatches 17; Indels 8; Gaps 5;

QY 1 TTTTCTTTTTCAGATGGAGTTTCCTCTCTGTCGCCAGGCTGGAGTGCATGCGCAA 60  
DB 8 TTTTCTTTTTCAGATGGAGTTTCCTCTCTGTCGCCAGGCTGGAGTGCATGCGCAA 67

QY 61 TCTAGCTCAGCGCAACTCCCGCTCCCGGGTTCAAGGATTCCTCTGCTCAGCTGCC 120  
DB 68 TCTAGCTCAGCGCAACTCCCGCTCCCGGGTTCAAGGATTCCTCTGCTCAGCTGCC 127

QY 121 CAGTAGCTGGGATACAGGCATGTCACCCAGCTCGGCTAATTTTGTATTTTTCAG 180  
DB 128 TAGTAGCTGGGATACAGGCATGTCACCCAGCTCGGCTAATTTTGTATTTTTCAG 187

QY 181 TAGAGATGGAGTTTCTCCATGTTGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC 240  
DB 188 TAGAGATGGAGTTTCTCCATGTTGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC 247

QY 241 CTCCTGCTCGGCTCCCAAGTGCT---AGATACAGGACTGGCCACCATGCCGG-CTC 296  
DB 248 CTCCTGCTCGGCTCCCAAGTGCTGAGATACAGGATGAGCCACCATGCCGGCTC 307

QY 297 TCCTGCTGCTAATTTTGTGTAGAACAGGTTTCACTGATG-TGCCCAAGCTGCTCTCC 355  
DB 308 TCCTGCTGCTAATTTTGTGTAGAACAGGTTTCACTGATGTTGCCAAGCTGCTCTCC 367

QY 356 TGAGCTCAGCAGTGCACCTGCTCAGCTCCCAAGTCTGGGATTCAGGCGCT--GCA 413  
DB 368 TGAGCTCAGCAGTGCACCTGCTCAGCTCCCAAGTCTGGGATTCAGGCGTGAGCC 427

QY 414 GCGTCTGCTGCTTTTATTTATTTTATTTTAAAGACAGGTGTCCTTACCCAG 473  
DB 428 ACCGTCTGCTGCTTTTATTTATTTTATTTTAAAGACAGGTGTCCTTACCCAG 487

QY 474 GATGAGTGCAGTGTGATCAGAGTCACTGCGGCTTCACTCTCAGATCAAGC-A 532  
DB 488 GATGAGTGCAGTGTGATCAGAGTCACTGCGGCTTCACTCTCAGATCAAGCAA 547

QY 533 TCCTCTGCTCAGCTCCCAAGTGTGGGACCAAGACATGACCACTACACCTGGCT 592  
DB 548 TCCTCTGCTCAGCTCCCAAGTGTGGGACCAAGACATGACCACTACACCTGGCT 607

QY 593 AATTTTATTTTAT 607  
DB 608 AATTTTATTTTAT 622

RESULT 2

BC068461/c  
LOCUS BC068461 3016 bp mRNA linear HTC 19-JUL-2004  
DEFINITION Homo sapiens cDNA clone IMAGE:30342539, with apparent retained intron.  
ACCESSION BC068461  
VERSION BC068461.1 GI:46249759  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS

1 (bases 1 to 3016)  
Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,K.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Faneb,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywicki,M.I., Skalek,U., Smalish,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932

2 (bases 1 to 3016)  
Strausberg,R.  
Direct Submission  
Submitted (02-APR-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Dr. Stefan Hansson  
cDNA library Preparation: Michael Brownstein / Ted Usdin  
Laboratory  
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAP Plate: 168 Row: m Column: 1  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21361917  
This clone has the following problem: retained intron.

FEATURES

source

1..3016  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30342539"  
/tissue\_type="Placenta, normal"  
/clone\_lib="NIH MGC\_147"  
/lab\_host="DH10B"  
/note="Vector: pBluescript"

ORIGIN

Query Match 31.9%; Score 460.6; DB 3; Length 3016;  
Best Local Similarity 66.8%; Pred. No. 3.5e-58;  
Matches 881; Conservative 0; Mismatches 344; Indels 94; Gaps 12;

QY 1 TTTTCTTTTTCAGATGGAGTTTTCCTCTCTGTTCCCGAGGCTGGAGTGCATGCGCAA 60  
DB 3012 TTTTCTTTTTCAGATGGAGTTTTCCTCTCTGTTCCCGAGGCTGGAGTGCATGCGCAA 2953

QY 61 TCTAGCTCAGCGCAACTCCGCTCCCGGGTTCAAGGATTCCTCTGCTCAGCTGCC 120  
DB 2952 TCTTGGCTCACTGCAACATCCGCTCCAGGGTTTCAAGCAATTTCTCTGCTCAGCTCCA 2893

Qy	121	CAGTAGCTGGGATTACAGGCATGTGCACCCAGCTCGGCTAAATTTTGTATTTTTTTTAG	180
Db	2892	GAGTAGCTGGGATTACAGGCATGTGCACCCAGCTCGGCTAAATTTG-----TGTFTTTAG	2837
Qy	181	TAGAGATGGAGTTTCTCATCTTGGTCAGGCTGTCTCGAACTCCGACCTCCGACCTCAGATCATC	240
Db	2836	TAGAGATGGGGTTTTGTCACTGTGGTCTGGCTGTCTTGAACCTGCTGACCTCAGGTGATC	2777
Qy	241	CTTCGGTCTCGGCTCCCAAAAGTGCTAGATACA---GGACTGCGCACATGCGCCGGC---	294
Db	2776	TGCCCGCTCGGCTCCCGAAGTCTCGGATTACAGGCATGAGCACCGTGCCTCGGCCTA	2717
Qy	295	-----TCTGCCCTGGC	304
Db	2716	GGTTCTTGCACTTTTGTGTAGGGCTAACAGTTTTCAAATAAACAACTTACTTTTGT	2657
Qy	305	TAAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGTGTCTCTGAGCTCAA	364
Db	2656	TTGTTTTGTGTGTAGACAGAGTTTTTTTTTGAGACACAGACCCAGGCTGGAGGGCAGTG	2597
Qy	365	GCAGTCCACCTG-----CCTACGCTCCCAAAGTGCTGGGATTA CAGGCGTGCAG	414
Db	2596	GCATGATCTCAGCTCACTGTAAACCTCCACCTCTCGGTTCAAGTGAATCTAGCGCCTCAG	2537
Qy	415	CCGTGCTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCCCACTTCTTACCCGG	474
Db	2536	CTCCTGAGAAGCTGGGATTACAGACATACACACCACACCTGGTTAAATTTTGTATTTT	2477
Qy	475	ATGAAGTGCAGTGTGTGATCA--CAGCTCACTGCAGCCTTCAACTCCTGAGATCAAG-C	531
Db	2476	TAGTAGAGACAGGATTTTCCACATGTGGCCAGGCTGTCTTGAACCTCTGGCCTCAAGTG	2417
Qy	532	ATCCTCTGCTCTAGCCTCCCAAGTAGCTGGGACAAAGAATGACACCACTACACTGTGC	591
Db	2416	ATCCACTCACCTCGGCTCCCAAAATTTCTGGGATTA CAGGCATGAGCCACACGCGCCGC	2357
Qy	592	TAAATTTTATTTTATTTTAAATTTTGTGAGACAGAGTCTCAACTGTGACACCCAGGCTG	651
Db	2356	CAACACAATTACT-TTTTTGTGTGTGTATGGAGTCTCG-CTCTGTCAACCAGGCTG	2299
Qy	652	GAGTGCAGTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGTTTCAAGTATTTCT	711
Db	2298	GAGTGCAGTGGCAAACTCGGCTCACTGCAACCTCGCCTCTCTGGGTTCAAGGATTTCT	2239
Qy	712	CTGCCCCAGCTCTCTGAGTAGCTGGGACTACAGGCGCCCAACACGCTAGCTAAAT-TTT	770
Db	2238	TCGTGCTCAGGCTCCTGAGTAGCTGGGAATACAGGCACACACCAACGCTCAGCTAAAT	2179
Qy	771	TTTGTAATTTTGTAGTAGAGA-TGGGGTTTCAACATGTTGCGCAGGTGATCTTGATCTCTGG	829
Db	2178	TTTGTAATTTTGTAGTAGACAGGGATTCACATGTAGGACGAGCTTCAATCTCTCTG	2119
Qy	830	ACCTTGATCTGCTGCTCGGCTCCCAAAGTGTGGGATTA CAGGCGTGAAGCAACCA	889
Db	2118	ACCTTGATCTCGCGGCTTGGCTTCCAAAGTGTGGGATGACAGGTGTGACCCACTG	2059
Qy	890	CGCCCCG-----TTATTTTAAATTTTGTGTGTTGTAAGAAATGCACTCA	934
Db	2058	CACCTGGCCCAACTTACTTTTAAAGTCAATCTTTTTTTTTTTTTTTTTTGTGCTGTTG	1999
Qy	935	CTCTGTTACCAAGCTGAGTGCAATGGGCAAAATCTGGCTCACTGCACTCTGCTCTCC	994
Db	1998	CTCTGTCAACCAAGCTGAGTGCAATTTTGGTCTCACTGCAACTTCAACTTCC	1939
Qy	995	CGGCTCAAGCGATTTCTCTGTCTCAGCTCCCAAGCAGCTGGGATTA CAGGCGCACTGCC	1054
Db	1938	CAGGTTCAAGCTATTTCTCTCGCTC--CATGTGAGTAGCTGGGATTA CAGGCTGCC	1881
Qy	1055	ACCACACCCGCTAAATTTTGTATTTTCAATAGAGCGGGTTTCAACATAATTTGTGACGG	1114
Db	1880	ACCATGCTGGCTAAATTTTGTATTTTGTAGATGGGGTTTCAACATAATTTGGCCAGG	1821

Qy	1115	CTGGTCTCAAACTCCTGACCTCAGGTGACCCACCTGGCTCAGGCTTCCAAAGTGTGGGA	1174
Db	1820	CTGGTCTCAAACTCCTGACCTCAGGTGACCCACCTGGCTCAGGCTTCCAAAGTGTGGGA	1761
Qy	1175	TTACAGGCGTGGAGCCACTCACCAGCGCGCTAATTAGATAAAAAAATATGTAGCAAT	1233
Db	1760	TTACAGGTGTAGCCACTGCGCTTGAAGTCACTTCTTATATATGTTCTTAACAAT	1702
RESULT 3			
BM996475			
LOCUS			
DEFINITION	UI-H-DT0-avl-i-16-0-UI.sl NCI CGAP_DTO Homo sapiens cDNA clone	549 bp mRNA linear	EST 17-JUN-2002
ACCESSION	IMAGE:5881047 3', mRNA sequence.		
VERSION	BM996475		
KEYWORDS	BM996475.1 GI:19721376		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutharia; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 549)		
TITLE	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
COMMENT	Unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a>		
	Tissue Procurement: Dr. Jose Mercuende		
	cDNA Library preparation: Dr. M. Bento Soares, University of Iowa		
	cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa		
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa		
	Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>		
	The following repetitive elements were found in this cDNA sequence: li-305, >ALU (matched complement) 221-440, >ALU (matched complement) 243-535, >SVA#Other		
	Seq primer: M13 FORWARD		
	POLYA=Yes.		

**FEATURES**  
**SOURCE**

FOIPA=rev.  
Location/Qualifiers

```

1..549
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5881047"
/tissue_type="Metastatic Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_DTO"
/note="Organ: Lung; Vector: pP73-Pac (P
modified polylinker; Site1: EcoR I; Site
NCI_CGAP_DTO is a cDNA library containi
tissue(s): Metastatic Chondrosarcoma in
was constructed according to Bonaldo, Le
Genome Research, 6:791-806, 1996. First
synthesis was primed with an oligo-dT pr
Not 1 site. Double stranded cDNA was lig
adaptor, digested with Not I, and cloned
into pP73-Pac vector. The oligonucleoti
the synthesis of first-strand cDNA conta
sequence that is located between the Not
(dT)18 tail. The sequence tag for this i
AACTGTCGG.

```

ORIGIN  
TAG\_SEQ=AACTGTTCCG<sup>18</sup>

Query Match 31.0%; Score 446.8; DB 5; Length 549;  
Best Local Similarity 95.4%; Pred. No. 6.8e-56;  
Matches 515; Conservative 0; Mismatches 17; Indels 8

QY 1 TTTT TTTT TGAGATGGAGTTTTCGCTCTTGTGCCCCAGGCTGGAGTGCAATGGCGCAA 60

```

Db      8 TTTT... 1842 bp
QY      61 TCTAGCTCAGCGCAACCTCCGCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCTCCC 120
Db      68 TCTAGCTCAGCGCAACCTCCGCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCTCCC 127
QY      121 CAGTACCTGGGATTACAGCATGTGACCCAGCTCCGCTAAATTTTATTTTATTTTAT 180
Db      128 TAGTACCTGGGATTACAGCATGTGACCCAGCTCCGCTAAATTTTATTTTATTT 187
QY      181 TAGAGATCGAGTTTCTCCATGTGCTCAGGCTGTGCTCGAATCTCCGACCTCAGATGATC 240
Db      198 TAGAGATCGAGTTTCTCCATGTGCTCAGGCTGTGCTCGAATCTCCGACCTCAGATGATC 247
QY      241 CTCCTGCTCCGCTCCCAAGTGTCT--AGATACAGGACTGGCCACCATGCCGCTC 296
Db      248 CTCCTGCTCCGCTCCCAAGTGTCTGAGATTACAGGCTGAGCCATGCCGCTC 307
QY      297 TGCCTGGCTTAATTTTGTGATGAACACAGGTTTCACTGATG-TGCCCAAGTGTCTC 355
Db      308 TGCCTGGCTTAATTTTGTGATGAACACAGGTTTCACTGATGTTGCCCAAGTGTCTC 367
QY      356 TGAGCTCAAGCAGTCCACTGCTCAGCTCCCAAGTGTGGGATTACAGGCTG--GCA 413
Db      368 TGAGCTCAAGCAGTCCACTGCTCAGCTCCCAAGTGTGGGATTACAGGCTGAGCC 427
QY      414 GCGTGGCTGGCTTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 473
Db      428 ACCGTGGCTGGCTTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 487
QY      474 GATGAGTGCAGTGTGATCACAGTCACTGACGCTTCACT--CCTGAGATCAAGCA 532
Db      488 GATGAGTGCAGTGTGATCACAGTCACTGACGCTTCACTCCCTGAGATCAAGCA 547

```

```

RESULT 4
BC011119
LOCUS   BC011119 Homo sapiens, clone IMAGE:3047997, mRNA.
DEFINITION BC011119
ACCESSION BC011119.1 GI:15029795
VERSION   HTCC
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 1842)
AUTHORS   Strausberg, R.
TITLE     Direct Submission
JOURNAL   Submitted (25-JUL-2001) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Baylor College of Medicine Human Genome
          Sequencing Center
          Center code: BCM-HGSC
          Web site: http://www.hgsc.bcm.tmc.edu/cdna/
          Contact: amg@bcm.tmc.edu
          Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louleaged, H.,
          Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
          A.N., Gibbs, R.A.
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
          Series: IRAK Plate: 19 Row: 0 Column: 1
          This clone has the following problem: retained intron.

```

```

FEATURES             source
Location/Qualifiers
1..1842
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3047997"
/tissue_type="Cervix, carcinoma"
/clone_lib="NIH_MGC_12"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"

ORIGIN
Query Match      27.5%; Score 396.2; DB 3; Length 1842;
Besc Local Similarity 70.3%; Pred. No. 1.1e-48;
Matches 746; Conservative 0; Mismatches 263; Indels 52; Gaps 14;

QY 2 TTTT... 61
Db 570 TTTT... 629
QY 62 CTCAGCTCACC... 121
Db 630 CTCGCTTACTGCAAGCTCCGCTCCCGGGTTCAACCTTTCTCTGCTCAGCCCTCCA 689
QY 122 AGTAGCTGGGATTACAGGCATGTGACCCACGCTCGGCTAAATTTTATTTTATTTTATTT 181
Db 690 AGTAGCTGGGACTACTGTGCTGGCCGACCAACGCTGGCTAAATTTT- TTTCTATTTTATTT 748
QY 182 AGAGATGAGATTCTCCATGTTGGTTCAGGCTGGTCTCGAATCTCCGACCTCAGATGATCC 241
Db 749 AGAGACGGGGTTTCAACCGTGTAGCCAGATGGTCTTGATGTCCTGACCT--TGTGATTC 806
QY 242 CTCGCTTCGGCTCCCAAAGTGTAGATACAGGACTGGCCACCATGTCGCCGCTCTGCGCT 301
Db 807 ACCTGCTCGGCTCCCAAAGTGTGGGATTACAGGAGTGAGCCACCGCCGCGCTTGA 866
QY 302 GGCTAAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTG-GTCTCTGAGC 360
Db 867 TTTTATTTTATTAATAGATGTTCTCACTATGTTGCCGAGCTAGTGTCTCAAACTCTCTGGGC 926
QY 361 TCAAGCAGTCCACCTGCTCAGCCCTCCCAAAGTGTGGGATTACAG--GCGTGACCGCT 418
Db 927 TCAAGCCATCTCCACCTT-GGCTCCCAAAGTGTGGGATTACAGTCTCAGCCACTGT 985
QY 419 GCTGGCTCTTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 478
Db 986 GCTGGCTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1038
QY 479 AGTGAGTGGTGTGATCACAGCTCACT-----GCAGCCTTCAAC 517
Db 1039 ACTGAGTGGTGGGATCTCAGCTCACTGCAACCTCGGCTAGTGTGAGCGCACTCTCGTC 1098
QY 518 TCCTGAGATCAAGC-ATCCTCTGCTCAGCTCCCAAAGTGTGGGACCAAGACATGC 576
Db 1099 TCCTGGGTTCAAGCAATTTCTCTGCTCAGCTCCCGAGTGTAGATTTACAGCGCAT 1158
QY 577 ACCACTACCTGGCTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 636
Db 1159 GCCACACGCTGGCTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1215
QY 637 CTGTCAACCCAGCTGGAGTGCAGTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCTCCG 696
Db 1216 CTGTCA--CAAGCTGGAGTGCAGTGGCGCTGATCTTGGTCTCACTGCAACCTCTGCTCTCCG 1274
QY 697 GGTTCAGATTATTTCTCTGCTCCCAAGCTCTCTGAGTGTGGGACTTACAGGCGCCACCCAC 756
Db 1275 GGTTCAGCCATTTCTCTG-GCCAGCTCTCAAGTGTGGGACTTACAGGCGCCGCGCAC 1333
QY 757 GCCTAGCTAAT-TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 814
Db 1334 CATGCCCGCTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1393
QY 815 GATCTTGATCTCTGGACCTTGTGATCTGCTGCTCGCTCGGCTCCCAAGTGTGGGATTAC 874

```



Further information about the clone and the sequencing project is available at <http://mips.qsf.de/projects/cdna/>.

**FEATURES**  
**source**

```

location/Qualifiers
1. .4191
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="R2P2:DKFZp666K083"
/db_xref="taxon:9606"
/clone="DKFZp666K083"
/tissue_type="stomach"
/clone_lib="666 (synonym: hsto2) . Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="genomic"
1. .4191
/gene="DKFZp666K083"
2517. .2675
/gene="DKFZp666K083"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH56307.1"
/db_xref="GI:52545758"
/translation="MMTROVSKTSHRSILFLVETGGCAASKHKGPGWVAHCNPSTL
GGGGGQIFR"

```

gene

gene	1. .4131
	/gene="DKFZp666K083"
CDS	2517. .2675

CDS 2517.2675

/gene="DKFZp666K083"

/codon start=1

/product="hypothetical protein"

```
/protein_id="CAH56307.1"
```

/db\_xref="GI:52545758"

```
/translation="MMTRQVSKTS
```

GGGGGIR"

## ORIGIN

Query Match	26.78;	Score 384.6;	DB 3;	Length 4191;
Best Local Similarity	62.08;	Pred. No. 4e-47;		
Matches 789;	Conservative 0;	Mismatches 399;	Indels 85;	Gaps 8;
Qy	1	TTTTTTTTTTTGAGATGGAGTGTTCGCTCTGTGTGCCACGGCTGGAGTGCAGATGGGGCAA	60	
Db	2905	TGTTTTGTTTTGAAGACCGGCTCTACGGCTGTCGCCCTAGGCTGGAGTGCAGGGTGGGA	2846	
Qy	61	TCTCAGTCTACCGGCAACCTCCGGCTCCCGGGTTCAAGCGATTCTCTGCCTCTACGCTCC	120	
Db	2845	TCTCGGCTCACTGCAACTCCGGCTCCCGGGTTCAAGCTATTCTCGTGCCTCAGCCTCC	2786	
Qy	121	CAGTAGCTGGGATTACAGGCGATGCAACCAACGCTCGGCTAAATTTTGTTATTTTTTTTAG	180	
Db	2785	AAGTAGCTGGAACCTACAGGGCGGTGCACATACACCCGCGTAAATTTTTCTGGCAATTTTAA	2726	
Qy	181	TAGAGATGGAGTTCCTCCATGTGGTTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC	240	
Db	2725	TAGAGACAGGGTTTCGCGCATGTTGGCCAGGCTGATCTCAAACTCCTGGCCTCATCTGATC	2666	
Qy	241	CCCTCGTCTCGGCTCCCAAGTGCTATAGATACA--GGACTGGCGCAACATGCCCGGCTCT	297	
Db	2665	TGCCCCCTCTGGCCCTCCCAAGTGCTGGGATTACAGCGGTGAGCCACATGCCCGGCTC	2606	
Qy	298	GCCTGGCTAAATTTTTGTGTGATAGAAACA--GGGTTCCTAGATGTGCCCAAGCTGGTCTCCT	356	
Db	2605	CCCTTAGTTTTGATGTCAGCACATCTCCAGTAATCTTAAGAACAGGCTTCCTGTGAGTT	2546	
Qy	357	GAGTCAAGCAGTCCACCTGCTCAGCTCCCAAGTGCTGGGATTACAGGCGTGAGCC	416	
Db	2545	GAGGTTTTTAGAGACCTGCTAGTCATCATGACAGGATCCTTTCAATGGGGGAGTCAATGT	2486	
Qy	417	GTGCCTGGCTTTTATTTTTATTTTTTAAAGACACAGGTGTCCCACTCTTACCCAGGAT	476	
Db	2485	CCATCAGTCAATATTCAAAATGTTCTTCTTTACTGCTTAATAACTTCTTCTACTCT	2426	
Qy	477	GAAGTCAGTGGTGTGATCAAGCTCACTGCAAGCTTCAACTCCTGAGATCAAGCATCT	536	
Db	2425	TCCCTTCA---TTCCTTTTTCTTGGGAGTCAGATGTTGAACTTCTCTGAAATGGTCACTCA	2369	
Qy	537	CCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGCACCACCTACACCTGGCTAATT	596	
Db	2368	ATTTCATTCCTCTTCTCAT-CTTTTGTTTATTAAACCTCTCTTTGTTTTTGAATATTCCT	2310	
Qy	597	TTTTATTTTTTATTTTTTATTTTTTTGAGACAGAGTCTCAACTCTGTGACCCAGGCTGGAGT	656	
Db	2309	TCTGGGACTTTTTTTTACTTTTTTTTTTTTGGACGGAGTCTCGCTCTATTGCCAGCGTG	2250	

Qy	657	CAGTGGCGCAATCTTTGGCTCAGCTGCAGCACTCTGCCTCCCGGGTTCAAGTTATTTCTCTCTGC	7116
Db	2249	AAGTGGCACAATCTGGGATCACTGCAAGCTTCAACCTCTGGGTTACGCGAATTTCTCCTAT	2190
Qy	717	CCAGCCTCTCTGAGTAGCTGGGACTACAGGCGGCCACCAAC- GCCTAGCTAAATTTTTTTTG	774
Db	2189	GTCHAGCCTCCGAGCAGCTGGGACCAAGGTGCCGCGCACCGCGCCGCCAATTTTTTG	2130
Qy	775	TATTTTTAGTAGAGA-TGGGTTTCAACATGTTCCGCAAGTTGATCTTGATCTCTGACACT	833
Db	2129	TATTTTTAGTAGAGACAGGGTTTCACTGTGTAGCCAGGATGGTCTCAATCTCTGACCT	2070
Qy	834	TGTGATCTGCTGCCTCGGCTCCCAAGTGTCTGGGATTACAGGCGTGAAGCCACCAAGCC	893
Db	2069	CGTGA-----TTGGCCTCCCAAGTGTCTGGGATTACAGGCAATGAGCCACAGCACC	2020
Qy	894	CGGCTTAT-----	901
Db	2019	TGGCCTTTTTTTTACTTTTATCTTTTCAACACCTCTATTGAAATTTTTTAAAGATATATAA	1960
Qy	902	-----TTTAAATTTTTTGTGTTTGAATCTCACTCTGTTTACCGAGCC	949
Db	1959	ATAACATAGTTGTTTTTTTGTCTTTTGTGACATACAGTCTCACTCTTTCAACCAAGC	1900
Qy	950	TGGAGTCAATGCCAAATCTCGGCTCACATGCAACCTCTGCCTCCCGGGCTCAAGCGATT	1009
Db	1899	TGGAGTCAAGTGTGTGATCTTGCTCAGTCTGCAACCTTCACTCCCGGGTTCAAGCAATT	1840
Qy	1010	CTCCTGCTCAGCCTCCCAAGCAGCTGGGATTACGGGCACCTCGCCACCAACACCCCGCTAA	1069
Db	1839	CTCCTGCTCAGCCTCCCAAGTAGCTGGGATTACAGGCGCTGCCACCAACCGCGCTAA	1780
Qy	1070	TTTTTGTATTTTCAATTAGAGCGGGGTTTCAACATATTTGTGAGGCTGTCTCAAACCTCC	1129
Db	1779	TTTTTGTATTTTGTAGTAGAGACGGGTTTCTCCATGTTGGCCAGGCTGTCTCCAACTCC	1720
Qy	1130	TGACCTCAGGTGACCCACCTGCTCAGCCTTCCAAAGTGTCTGGGATTACAGGCGGTGAGCC	1189
Db	1719	TGACCTCAAGTGATCCGCCACACTTGGCCTCCCAAAAGTGTCTGGGATTACAGGTGTGAGCC	1660
Qy	1190	ACCTCACCCAGCC	1202
Db	1659	ACCGCACTGGCC	1647

RESULT 7			
AV762220/c	AV762220	1345 bp	mRNA linear EST 19-OCT-2000
LOCUS	AV762220	MDS Homo sapiens cDNA clone	MDSCBD07 5', mRNA sequence.
DEFINITION	AV762220		
ACCESSION	AV762220		
VERSION	AV762220.1	GI:10920068	
KEYWORDS	EST.		

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 1345)

**AUTHORS**  
Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,  
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,  
Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M.,  
Lu, G., Yang, Y., Gao, G., Zhang, Q., Chen, S., Han, Z. and Chen, Z.

**TITLE**  
Homo sapiens cDNA MDS clones

Unpublished (2000)  
Contact: Zequang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919 (ex.45)  
Fax: 86-21-50801922

**FEATURES**  
Location/Qualifiers  
This clone is available at CHGC in Shanghai.  
Email: hanzg@chgc.sh.cn



```

source
1. 1345
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MDSCBD07"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/clone_lib="MD5"
/notes="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"

ORIGIN
Query Match      26.6%; Score 383.6; DB 1; Length 1345;
Best Local Similarity 82.8%; Pred. No. 8.4e-47;
Matches 502; Conservative 0; Mismatches 86; Indels 18; Gaps 5;

QY 600 ATTATTTATTTTAAATTTTGTGACAGAGTCTCAACTCTGTCTACCCAGGCTGAGTGCAG 659
Db 606 ATTTTNTNTNTTTTGTGACAGGAGTCTC-ACTCTGTCTGCCAGGCTGAAGTGCAG 548
QY 660 TGGCGCAATCTGGCTCACTGCACCACTCTGCCTCCCGGGTTCAAGTTATTCTCTGCCCC 719
Db 547 TGGGGTCACTCAGCTCACTGCACCTCTGCCTCTGGGTTCAAGCAATCTCTGCCCTC 488
QY 720 AGCTCTCTAGTAGTGGGACTACAGCGC---CCACACGCTAGCTAATTTTTTGTGA 776
Db 487 AGCTCTCTAGTAGTGGGACTACAGCGCGCTCCACACGCTAGCTAATTTTTTGTATT 428
QY 777 TTTTGTAGTAGATGGGTTTCCATGTTTGGCAGGCTTGTCTTGAACCTCTGACCTCGT 836
Db 427 TTTAGTAGAGACA-GGGTTTTCATGTTTGGCAGGCTTGTCTTGAACCTCTGACCTCGT 369
QY 837 GATCTGCTGCTCGGCTCCCAAGTCTGCGGATACAGGCGTGGAGCCACACACCGCCG 896
Db 368 GATCCACCCACACGAGTCTCTCAAGTCTGGGATACAGGCGTGGAGCCACACACCGCCG 309
QY 897 C-----TTATTTTAAATTTTGTGTAATTTGAAATCTCACTCTGTGTACCC 945
Db 308 CTTCAATTTGCCATTTCTTTTTTTTTTTTCTTGTGAGATGGAGTCTCACTCTATTGCC 249
QY 946 AGCTGGAGTGCATGGCCAAATCTGGCTCACTCACTCACTCTGCTCCGGCTCAAGC 1005
Db 248 AGCTGGAGTGTGGGCAATCTGGCTCACTCACTCACTCTGCTCCGGATTCAGT 189
QY 1006 GATTCTCTGTCTCAGCTCCCAAGCAGTGGGATACAGGCGCTGCAACACACACCGC 1065
Db 188 GATTCTCTGTCTCAGCTCCCAAGCAGTGGGATACAGGCGCTGCAACACACCGC 129
QY 1066 CTAATTTTGTATTTTCAATAGAGCGGGTTTCAACATATTTGTGAGGCTGTCTCAAA 1125
Db 128 CTAATTTTGTATTTTCAATAGAGCAGGGTTTCAACATATTTGTGAGGCTGTCTCAAA 69
QY 1126 CTCTGACCTCAGTGCACCTGCTGAGCTTCCAAAGTCTGGGATACAGCGCTG 1185
Db 68 CTCTGACCTC--GTATCGCCCGCTCAGCTTCCAAAGTCTGGGATACAGCGCTG 11
QY 1186 AGCCAC 1191
Db 10 AACAC 5

RESULT 8
BQ722917/c
LOCUS
DEFINITION
BQ722917
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BQ722917 910 bp mRNA linear EST 16-JUL-2002
IMAGE:6190443 5', mRNA sequence.
BQ722917.1 GI:21861803
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 910)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Prepared by: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13589 row: e column: 04
High quality sequence stop: 653.
Location/Qualifiers
1. 910
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6190443"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/clone_lib="Lupski sympathetic trunk"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGCTCG-3' and
5'-GACTAGTCTAGTTCGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

ORIGIN
Query Match      26.3%; Score 379; DB 5; Length 910;
Best Local Similarity 80.8%; Pred. No. 4.5e-46;
Matches 514; Conservative 0; Mismatches 91; Indels 31; Gaps 5;

QY 595 TTTTATTTTATTTTATTTTGTGAGACAGAGTCTCAACTCTGTCAACCCAGGCTGGAG 654
Db 643 TTTTATTTTATTTTATTTTGTGAGCGGAGTCT-TGCTCTGTGCGCCAGGCTGGAG 595
QY 655 TGCAGTGGCGCAATCTGGCTCACTGCACCTCTGCCTCCCGGGTTCAAGTTATTCTCT 714
Db 584 TGCAGTGGCGCAATCTGGCTCACTGCACCTCTGCCTCCAGGTTCAAGCGATTCTCT 525
QY 715 GCCCGAGCTCTGAGTAGCTGGGACTACAGGCG---CCACACGCTAGCTAATTTT 771
Db 524 GCTCAGCTCTCTGAGTAGCTGGGACTACAGGCAATGCCACGCTCGGCTAATTTT 465
QY 772 TTTGATTTTGTAGAGATGGGGTTTCAACATGTTGCGCAGGTTGATCTTGTGAC 831
Db 464 TGTATTTTATGGAGACAGGGTTTCAACGCTGTAGCCAGGATGGTCTGATCTCTGAT 405
QY 832 CTTGTGATCTGCTGCTCGGCTCCCAAAGTGTGGGATTCAGCGTGGAGCCACG 891
Db 404 CTTGTGATCTGCTGCTTGGCTCCCAAAGTGTGGGATTCAGCGTGGAGCCACGCG 345
QY 892 CCGGCG-----TTATTTTAAATTTTGTGTTTGTGAATCG 927
Db 344 CCGGCGCTCCCTAGTATTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGACGCG 285
QY 928 AATCTCACTCTGTATCCAGGCTGGAGTGCATATGGCCAAATCTCGGCTCACTGCAACCTC 987
Db 284 AGCTGCTCTGTCTGCGCCAGGCTGGAGTGCAGTGGCGGATCTCGGCTCACTGCAAGCTC 225
QY 988 TGCCTCCCGGGCTCAAGCGATTCTCTGTCTCAGCTCCCAAGCAGCTGGGATTCAGGCG 1047
Db 224 TGCCTTCGGGTTCCAGCGATTCTCTGTCTCAGCTCCCAAGTAGCTGGGACTACAGCG 165

```

```

Qy 1048 ACCTGCCACACACCCCGCTAA-TTTTGTATTTTCAATTAGAGCGGGGTTTCCACATAT 1106
Db 164 GCCCGCCATAGCCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCCACCGTTT 105
Qy 1107 TTGTGAGGCTGTCTCAAACTCTGACCTCAGGTGACCCACCTGCTCAGCCCTCCAAAG 1166
Db 104 TAGCCGGGATGTCTCGATCTCTGACCTC--GTGATCGCCCGCTCGGCCCTCCAAAG 47
Qy 1167 TGCTGGGATTACAGCGGTGAGCCACCTCACCAGCG 1202
Db 46 TGCTGGGATTACAGCGGTGAGCCACCGCGCCCGGCC 11

RESULT 9
AV700498 LOCUS 767 bp mRNA linear EST 16-JAN-2002
DEFINITION AV700498 GKC Homo sapiens cDNA clone GKCAAC05 3', mRNA sequence.
ACCESSION AV700498
VERSION AV700498.1 GI:10302469
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 767)
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE 21625106
PubMed 11752456
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1..767
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GKCAAC05"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GKC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Query Match 26.2%; Score 378.4; DB 1; Length 767;
Best Local Similarity 83.1%; Pred. NO. 5.9e-46;
Matches 481; Conservative 0; Mismatches 87; Indels 11; Gaps 4;

Qy 635 CTCTGTCAACCCAGGTGGAGTCAGTCGGCAATCTTGGCTCACTGCACCTTCGCTCC 694
Db 13 CTCTGTCAACCCAGGTGGAGTCAGTCGGCAATCTTGGCTCACTGCACCTTCGCTCC 72
Qy 695 CGGGTTCAAGTTATTTCTCTCGCCCGCCAGCTCTCTGAGTAGCTGGGACTACAGCG--CCC 751
Db 73 CGGGTTCAACCCATTTCTCTGCTCAACCTCCCGAGTAGCTGTGAATACAGCGGTCGCGC 132
Qy 752 ACCAGCCTAGCTAATTTTTTTGTATTTTTTAGTAGAGATGGGG-TTCACCATGTTCGCCA 810
Db 133 ACCAGCCCGCACTAATTTTTTTGTATTTTTTAGTAGAGACGGGGTTTACCGGTGTAGCCA 192

```

```

Qy 811 GGTTCATCTTGTATCTCTGAGCTTGTGATCTGCCTGCTGCTCCGCTCCCAAGTGTGGGA 870
Db 193 GGATGGTCTGGATCTCTGAGCTGCTGATCTGCCTGCTGCTCCGCTCCCAAGTGTGGGA 252
Qy 871 TTACAGGCGTGAAGCCACACCGCCGGCTTATTTTAAATTTTTTGTGTTGAAATGGAAT 930
Db 253 TTACAGGCTTGAGCCACCGCAGCCCGCCTCT-----TATTTTTTTTGTGATGGAGT 306
Qy 931 CTCACCTGTATCCAGGCTGGAGTGCAAATGGCCAAATCTCGGCTCACTGCACACCTCTGC 990
Db 307 CTCACACTGTCACTGGGCTGGAGTGCAGTGGAGCGATCTCGGCTCACTGCACACCTCCGC 366
Qy 991 CTCGCGGCTCAAGCGATTCTCTGCTCAGCTCCCAAGCAGCTGGGATACGGGCACC 1050
Db 367 CTCCTGGGTTCAAGAGATTCTCTGCTCAGCTCCCAAGTAGCTGGGATACAGGTGCC 426
Qy 1051 TGCCACCAACACCCCGCTA-ATTTTGTATTTTCAATAGAGCGGGTTTCCACATATTG 1109
Db 427 CACCAACACGCTGGCTAGTTTTTTGTATTTTGTAAAGATGGGTTTCCACATGCTGG 486
Qy 1110 TAGGCTGGTCTCAAACTCTGACCTCAGGTGACCCACCTGCTCAGCTTCCAAAGTGC 1169
Db 487 CCAGGCTGGTCTTGAATCTCTGACATCAGGTGATCCGNCACCTTAGCCCTCCCAAGTGC 546
Qy 1170 TGGGATTACAGCGGTGAGCCACCTCACCAGCGCGCTAA 1208
Db 547 TTGGATTACAGCGGTGAGCCACCTACCTGGCCAGCAAA 585

RESULT 10
CD709038 LOCUS 613 bp mRNA linear EST 25-JUN-2003
DEFINITION EST25565 human nasopharynx Homo sapiens cDNA, mRNA sequence.
ACCESSION CD709038
VERSION CD709038.1 GI:32239668
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 613)
AUTHORS Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and
Zeng,Y.-X.
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx
JOURNAL Unpublished (2003)
COMMENT Contact: Yixin Zeng
Cancer Center
Sun Yat-sen University
651 Dongfeng Road East, Guangzhou 510060, China
Tel: 86-1380-9770-743
Fax: 86-20-8775-4506
Email: yxzeng@gzsums.edu.cn.
Location/Qualifiers
1..613
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"

FEATURES
source
1..613
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="normal nasopharynx"
/clone_lib="human nasopharynx"
/note="ESTs generated from a normal nasopharynx cDNA
library from southern Chinese"

ORIGIN
Query Match 26.2%; Score 377.6; DB 6; Length 613;
Best Local Similarity 83.0%; Pred. NO. 8.4e-46;
Matches 483; Conservative 0; Mismatches 84; Indels 15; Gaps 4;

Qy 634 ACTCTGTCAACCCAGGTGGAGTGCAGTCGGCAATCTTGGCTCACTGCACCTTCGCTCC 693
Db 610 ACTCTGTCCGCGCTGGAGTGCAGTCGGCGTGATCTCAGCTCACTGCACACCTCTGCCTC 551
Qy 694 CCGGGTTCAAGTTATTTCTCTCGCCCGCCAGCTCTCTGAGTAGCTGGGACTACAGGCGCCAC 753

```

Db 550 CTGGGTTCAAGCAATCTCTCTGCTCAGCCTCTGAGTAGCTGGGACTACAGCGCCGC 491  
Qy 754 CAGCCTAGCTAAT-TTTTTTGTATTTTAGTAGAGA-TGGGGTTCCACCATGTTGCCAG 811  
Db 490 CACCAGCCCGCAATATTTTGTATTTTAGTAGAGACAGGGTTTCATCATGTTGCCAG 431  
Qy 812 GTTGATCTTGATCTCTGGACCTTGATCTGCTGCTCGGCTCCCAAAGTCTGGGAT 871  
Db 430 GCTTGTCTTGAATCTCTGACCTCGTGATCCACCCACCGACTCTCAAGTCTGGGAT 371  
Qy 872 TACAGCGGTGAGCCACACGCCCGGC-----TTATTTTAAATTTTGTGTTT 920  
Db 370 TACCGCGGTGAGCCACCATGCGCGGCTCATTTGCGCAATCTTTTTTTTTTCTTT 311  
Qy 921 GAAATGGAATCTCAGCTCTGTATCCAGGCTGGAGTGCAATGGCCAAATCTCGGCTCACATG 980  
Db 310 GAGATGGAGTCTCAGCTCTATTTGCCAGGCTGGAGTCTGTGGACAAATCTCGGCTCACATG 251  
Qy 981 CAACCTCTGCTCCCGGCTCAAGCGATTTCTCTGTCTCAGGCTCCCAAAGCTGGGAT 1040  
Db 250 CAACCTCTGCTCCCGGATCAAGTGATTTCTCTGCTCAGGCTCCGAGTAGCTGGGAT 191  
Qy 1041 TACGGCAGCTGCCACACACCCGCTAATTTTGTATTTTCAATTAGAGCGGGTTTCA 1100  
Db 190 TACAGGCATGCCACCATGCGCAGCTAATTTTGTATTTTACTAGAGACAGGGTTTCA 131  
Qy 1101 CCATATTTCTAGGCTGGTCTCAAACTCTGAGCTCAGGTGACCCACCTGCTCAGGCTT 1160  
Db 130 CCATATTTGTAGGCTGGTCTCAAACTCTGAGCTC--GTGATCGCGCCGCTCAGGCTC 73  
Qy 1161 CCAAGTGTGGGATTTACAGGCTGAGCCACCTCACCCAGCC 1202  
Db 72 CCAAGTGTGGGATTTACAGGCTGAGCCACCTCACCCGCGCC 31

## RESULT 11

CR614786  
LOCUS  
DEFINITION  
full-length cDNA clone CS0DF06YD23 of Fetal brain of Homo sapiens (human)

ACCESSION  
CR614786.1 GI:50495593

VERSION  
HTC; CDS; CDS; CDS

KEYWORDS  
Homo sapiens (human)

SOURCE  
Homo sapiens

ORGANISM  
Homo sapiens

## REFERENCE

AUTHORS  
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE  
Full-length cDNA libraries and normalization  
JOURNAL  
Unpublished  
REMARK  
Contact: Feng Liang Email: fliang@lifetech.com URL:  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue  
Genoscope.  
Direct Submission  
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

## REFERENCE

AUTHORS  
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE  
Full-length cDNA libraries and normalization  
JOURNAL  
Unpublished  
REMARK  
Contact: Feng Liang Email: fliang@lifetech.com URL:  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue  
Genoscope.  
Direct Submission  
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

## COMMENT

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
Location/Qualifiers  
1. .2097  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DF06YD23"  
/tissue\_type="Fetal brain"

## FEATURES

source

## RESULT 12

CA439577

LOCUS

DEFINITION

UI-H-D10-aux-p-06-0-UI.s1

NCI CGAP D10 Homo sapiens

CDNA clone

linear

EST 08-NOV-2002

604 bp

mRNA

/plasmid="pCMVSPORT\_6"

## ORIGIN

Query Match 26.2%; Score 377.6; DB 3; Length 2097;  
Best Local Similarity 71.6%; Pred. No. 5.4e-46;  
Matches 614; Conservative 0; Mismatches 219; Indels 25; Gaps 8;  
Qy 352 CTCCTGAGCTCAAGCAGTCCACCTGCTCAGCCTCCCAAGTGTGGGATTAACAGCG-- 409  
Db 480 CTCCTGGGTTCAAGCTATTCTCTGCTCAGCTTCTGAGTAGCTGGGATTAATAGCGCA 539  
Qy 410 TGCAGCGCTGCTGCTGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 469  
Db 540 TGCCACCATGCTGCTGCTAATTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 599  
Qy 470 CCAGGATGAAGTGCAGTGTGTGATCAGACCTCAATGATCCACCTGCTCTCTCTCTCTCTCTCT 529  
Db 600 CAGGCTGGTCTTGAATCTCTGACCTCAATGATCCACCTGCTCTCTCTCTCTCTCTCTCTCT 659  
Qy 530 GCATCTCTCTGCTCAGCCTCCCAAGTGTGGGATTAACAGCAGTGTGGGATTAACAGCGCT 589  
Db 660 GGGATTAGAGGTATGAGCCACCATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 712  
Qy 590 GCTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 649  
Db 713 -----TTAATTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 765  
Qy 650 TGGAGTGCAGTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTAT 709  
Db 766 TGGAGTGCAGTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTAT 825  
Qy 710 CTCTGCTCCCGGCTCTGAGTAGTGGGATTAACAGCAGTGTGGGATTAACAGCGCTAGCTAA 768  
Db 826 CTCTGCTCCCGGCTCTGAGTAGTGGGATTAACAGCAGTGTGGGATTAACAGCGCTAGCTAA 885  
Qy 769 TTTTGTATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 827  
Db 886 TTTTGTATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 945  
Qy 828 GGAACCT--TGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 885  
Db 946 TGACCTCAGGTGATTCACCCCGGCTCTCCCAAGTGTGGGATTAACAGCGCTAGCTAA 1005  
Qy 886 ACCAGCGCGGCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 945  
Db 1006 ACTGCGCTGCGCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 1061  
Qy 946 AGGCTGGAGTGAATGGCCAAATCTGCGCTCACTGCAACCTCTGCTCTCCCGGGCTCAAGC 1005  
Db 1062 AGGCTGGAGTGAATGGCCAAATCTGCGCTCACTGCAACCTCTGCTCTCCCGGGCTCAAGC 1121  
Qy 1006 GATTCCTCTGCTCA--GCTCCAGCAGCTGGGATTAACAGCAGTGTGGGATTAACAGCGCTAG 1064  
Db 1122 GATTCCTCAGGCTCAGGCTCTCCCAAGTGTGGGATTAACAGCGCTAGCTAA 1181  
Qy 1065 GCTAATTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1124  
Db 1182 GCTAATTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1241  
Qy 1125 ACTCCTGACCTCAGGTGACCCACCTGCTCAGCCTTCCAAAGTGTGGGATTAACAGCGCT 1184  
Db 1242 ACTCCTGACCTCAGGTGACCCACCTGCTCAGCCTTCCAAAGTGTGGGATTAACAGCGCT 1301  
Qy 1185 GAGCCACCTCACCCAGCC 1202  
Db 1302 GAACCATTCGCCAGCC 1319

## RESULT 12

CA439577

LOCUS

DEFINITION

UI-H-D10-aux-p-06-0-UI.s1

NCI CGAP D10 Homo sapiens

CDNA clone

linear

EST 08-NOV-2002

604 bp

mRNA

ACCESSION CA439577  
 VERSION CA439577.1 GI:24803997  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 604)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Dr. Jose Mercuende  
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu  
 The following repetitive elements were found in this cDNA sequence: 11-298, >ALU (matched complement) 302-590, >ALU (matched complement)  
 Seq primer: M13 FORWARD  
 POLYA=Yes

FEATURES  
 source  
 Location/Qualifiers  
 1..604  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-H-D10-aux-p-06-0-UI"  
 /tissue\_type="Lung Focal Fibrosis"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI CGAP D10"  
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI-CGAP D10 is a cDNA library containing the following tissue(s): A pool of lung Focal Fibrosis. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ATACGCGTTC.  
 TAG TISSUE=lung with fibrosis  
 TAG\_LIB=UI-H-D10  
 TAG\_SEQ=ATACGCGTTC

ORIGIN  
 Query Match 26.2%; Score 377.4; DB 6; Length 604;  
 Best Local Similarity 81.1%; Pred. No. 9e-46; Mismatches 101; Indels 14; Gaps 4;  
 Matches 492; Conservative 0;

QY 602 TTTTATTTTAAATTTTTCAGACAGAGTCTCAACTCTGTCCACCCAGGCTGGAGTGCAGTG 661  
 Db 1 TTTTATTTTATTTTTCAGATGGAGTCTC-ACCTGTGTCACCCAGGCTGGAGTGGGTG 59  
 QY 662 GGCATATCTGGCTCACTGCAACCTCTGCCTCCGGGTTCAAGTTATTCCTCTGCCCGCAG 721  
 Db 60 GTGCAATATCGGCTCACTGCAACTTCTGCTCCAGGTTTCAAGCAATTCCTCGCTCAG 119  
 QY 722 CCTCTGAGTAGTGGAGCTACAGGCGCCACACGCTAGCTAAT-TTTTGTGATTTT 780  
 Db 120 CTTCCGGAGTAGTGGAGCTACAGGCGCCACGCGCCACACACCGAGTAAATTTTGTATTT 179  
 QY 781 TAGTAGAGATGGGG-TTCACCATGTTGCGCAGGTTGATCTTGATCTCTGACCTTGTGAT 839  
 Db 180 TAGTAGAGATGGGTTTACCATGTTGGCCAGGATGGTCTTGATCTCTTGACCTTGTATT 239

QY 840 CTGCTGCTCGGCTCTCCAAAGTGTGAGATTACAGGCGTGAGCCACACGCGCGGCTT 899  
 Db 240 CTGCCCACTCAGCCCTCCCAAAGTACTAGGATTACAGGCATGAGCCACCATGCTGCGC-- 297  
 QY 900 ATTTTAAATTTTGTGTTTGTGAAATGAATCACTCTCTGTACCCAGGCTGGAGTGCAA 959  
 Db 298 -----CTTTTATTTTGTGAGACAGAGTCTCGCTCAGTTGCCAGGCTGAGTGCGAG 348  
 QY 960 TGGCCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCTGTCTC 1019  
 Db 349 TGGCGTGAATTCAGCTCACTGCAATCTCCACTCCAGGTTCAAGCAATTCCTCGTGCCTC 408  
 QY 1020 AGCTTCCCAAGCAGCTGGGATTACGGGCACCTGCCACACACCCCGCTAAATTTTGTATT 1079  
 Db 409 AACCTCCCGAGTAGTTGAGACTACAGGCGCCCGCCACACACACCGAGCTGATTTTTTTT 468  
 QY 1080 TTCAATTAGAGCGGGGTTTCAACATATTTGTTCAGGCTGTCTCAAACTCCCTGACCTCAGG 1139  
 Db 469 TTTAGTAGACAGAGGTTTCAACATATTTGGGAGGCTAGTCTTGAATCTCTGACCTCAGG 528  
 QY 1140 TGACCCACTCTGCTCAGCCTTCCAAAGTGTGGGATTACAGCGGTGAGCCACCTCACCCA 1199  
 Db 529 TGATCCACTCTGCTCGGCTCCCAAAGTGTGAGATTACAGGTGTAGCCATGCGCGCTG 588  
 QY 1200 GCGGCT 1206  
 Db 589 GCAACT 595

RESULT 13  
 LOCUS CD690152/c  
 DEFINITION EST6675 human nasopharynx Homo sapiens cDNA, mRNA sequence.  
 ACCESSION CD690152  
 VERSION CD690152.1 GI:32210629  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 628)  
 REFERENCE Liu,X.-Q., Zhou,Y., Zhang,L.-J., Xu,H., Chen,H.-K., Pan,Z.-G. and Zeng,Y.-X.  
 Transcriptional Gene Expression Profile of Human Nasopharynx  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Yixin Zeng  
 Cancer Center  
 Sun Yat-sen University  
 651 Dongfeng Road East, Guangzhou 510060, China  
 Tel: 86-1380-9770-743  
 Fax: 86-20-8775-4506  
 Email: yxzeng@gzsums.edu.cn.

FEATURES  
 source  
 Location/Qualifiers  
 1..628  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="normal nasopharynx"  
 /clone\_lib="human nasopharynx"  
 /note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

ORIGIN  
 Query Match 26.1%; Score 376.4; DB 6; Length 628;  
 Best Local Similarity 81.3%; Pred. No. 1.2e-45;  
 Matches 491; Conservative 0; Mismatches 96; Indels 17; Gaps 4;

QY 614 TTTTGTAGACAGAGTCTCAACTCTGTCCAGGCTGGAGTGCAGTGGCGCAATCTTGG 673  
 Db 625 TTTTGTAGACGCGAGTCTCACTCTGTCCCAAGGCTGGAGTGCAGTGGCGTATCTCAG 566  
 QY 674 CTCACCTGCAACCTCTGCTCGGCTTCAAGTTATTTCTCTGCCCGCAGCCTCTCTGAGTAG 733



Wiemann, S.  
The German cDNA Consortium  
Direct Submission  
Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuerberg, GERMANY  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by EMBL (European Molecular Biology Laboratories,  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
This clone (DKFZp5471147) is available at the RZPD Deutsches  
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.  
Please contact RZPD for ordering:  
<http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp5471147>  
Further information about the clone and the sequencing project is  
available at <http://mips.gsf.de/projects/cdna/>.

# FEATURES

source

1. .2777  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp5471147"  
/tissue\_type="brain"  
/clone\_lib="547 (synonym: hfbrl). Vector pSport1; host  
DH10B; sites NotI + Sali"  
/dev\_stage="fetal"  
/note="genomic"

# ORIGIN

Query Match 26.0%; Score 375.4; DB 3; Length 2777;  
Best Local Similarity 71.3%; Pred. No. 1e-45;  
Matches 581; Conservative 0; Mismatches 221; Indels 13; Gaps 6;  
  
QY 605 TATTTTAAATTTTGTAGACAGAGTCTCAACTCTGTACCCAGGCTGGAGTGCAGTGGCG 664  
Db 2777 TTTTATTTTATTTTGTAGATGGAGTCTC-GCTCTGTGCTCAGGCTGGAGTGCATGGCG 2719  
  
QY 665 CAATCTTGGCTCACTGCACCTCTGCTCCGGTTCAGTTATTTCTCTGCTGCCAGCT 724  
Db 2718 CAATCTTGGCTCACCACAGCTCCGGCTCTCGGTTTAAAGCATTTCTCTGCTCAGCT 2659  
  
QY 725 CTGTAGTAGCTGGACTACAGGC--GCCACACAGCTAGCTAAATTTTGTATTTT 781  
Db 2658 CCGAATAGCTGGACTATAGCATGTGCCCCACCCAGCTAATTTTGTATTTTAG 2599  
  
QY 782 AGTAGAGATGGGTTTCCATGTTCCGAGGTTGATCTTGTATCTCTGGACCTTTGATCT 841  
Db 2598 TAGAGACA-GGGTTTCCATGTTGGCCAGGATGGTCTCGATCTCTTAACCTTTGTATCC 2540  
  
QY 842 GCTGCTCGGCTCCCAAGTGTGGGATTACGCGTGAGCCACCCAGCCCGGCTTAT 901  
Db 2539 GCCCGCTCAGGCTCCCAAGTGTGGGATTACGAGTGAGCCACCGTGGC--CT 2482  
  
QY 902 TTTTAAATTTTGTGTTGTTGAAATGGAATCTCACTCTGTACCCAGGCTGGAGTGCATG 961  
Db 2481 TCTTATTTATTTATTTTGTAGGTGGAGTCTCACTCTGTGCTCAGGCTGGTGTGCAATG 2422  
  
QY 962 GCCAAATCTGGCTCACTGCAACCTTGTGCTCCCGGGCTCAAGCGATTCTCTGTCTCAG 1021  
Db 2421 ATGCAATCTGGCTCACTGCAACCTTGTGCTCCCGGGCTCAAGCGATTCTCTGTCTCAG 2362  
  
QY 1022 CTTCCCAAGCAGTGGGATTACGGCACCTGCGCACCAACCCGCTAATTTTT--GTATT 1079  
Db 2361 CTTCTGTAGTAGCTGAGATTACAGGGATGTGCCATGTGACCGCTGGCTAATTTTGGTATT 2302  
  
QY 1080 TTTCAATTAGAGGGGGTTTTCACCATATTTGTGAGGCTGGTCTCAAACTCTGACCTCAGG 1139  
Db 2301 TTTAGTAGAGATGGGTTTTCGCCACGTTGACCGGGCTGGTCTGCAACTCTGGCTCAG 2242  
  
QY 1140 TGACCCACTGCTCAGCTTCCAAAGTGTGGGATTACAGGCTGAGCCACTCACCCA 1199  
Db 2241 TGATTGCGCCGCTCAGCTCCCAAGTGTGGGATTACAGGATGAGCCACACACCTG 2182

QY 1200 GCCGCTAAATTTAGATAAAAAAATATGTAGCA----ATGGGGGGTCTTGTCTATGTTGCC 1255  
Db 2181 GCCTTTTCTTTTCTTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2122  
  
QY 1256 AGGCTGGTCTCAAACTTCTGGCTTCATGCAATCTCTTCCAAATGAGGCCAACACCCAGCC 1315  
Db 2121 AGGTTGATGTGCAAGTGGGCAATCTTGGCTCACCACCGCTCTCTCTCTCCAGGTTCC 2062  
  
QY 1316 AGTCACATTTTAAACAGTTACATCTTTATTTTAGTATATCTAGAAAGTAAATACATAAA 1375  
Db 2061 AGTGATTTCTGGCTCCAGCTCCCAAGTAGCTGGAGTACAGGATGTGCCACCACACC 2002  
  
QY 1376 CATGTCAAAACCTGCAGAAATTCAGTAGTAACAGAGTT 1410  
Db 2001 CAGCTAATTTTGTATTTTATTTAGTACAGAGGTT 1967

Search completed: September 15, 2005, 21:58:53  
Job time : 5124 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 14:42:12 ; Search time 6434 Seconds  
(without alignments)  
10859.873 Million cell updates/sec

Title: US-09-380-203-1  
Perfect score: 1442  
Sequence: 1 TTTT...TTTGGAGGAG.....TTAAACAAGCTTTAGACA 1442

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htgi.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1442	100.0	1442	6	AR051550 Sequence
2	1442	100.0	1442	6	AR072690 Sequence
3	1442	100.0	1442	6	AR073235 Sequence
4	1442	100.0	1442	6	AX594290 Sequence
5	1442	100.0	1442	6	AX599994 Sequence
6	1442	100.0	1442	6	AX709024 Sequence
7	1442	100.0	1442	6	AX709356 Sequence
8	1442	100.0	1442	6	AX770807 Sequence
9	1442	100.0	1442	6	AX774635 Sequence
10	1442	100.0	1442	6	BD070526 Transgeni
11	1442	100.0	1442	9	AF010144 Homo sapi
12	1301.4	90.2	124001	9	HS886822 Human DNA
13	1223.4	84.8	1418	6	BD070528 Transgeni
14	1080.2	74.9	1381	6	AR051479 Transgeni
15	1080.2	74.9	1381	6	AR072619 Sequence
16	1080.2	74.9	1381	6	AR073164 Sequence
17	1080.2	74.9	1381	6	BD070527 Transgeni
18	592.2	41.1	161074	2	AC087434 Pan trogl
19	592.2	41.1	162240	2	AC145756 Pan trogl

20	590.4	40.9	99370	9	AC005057	AC005057 Homo sapi
21	587.8	40.8	65608	6	AX330738	AX330738 Sequence
22	587.8	40.8	65608	6	AX332242	AX332242 Sequence
23	587.8	40.8	65608	6	AX335496	AX335496 Sequence
24	587.8	40.8	65608	9	HSU62293	U62293 Human LIM-k
25	587.8	40.8	67046	9	HSU63721	U63721 Human elast
26	587.8	40.8	148525	2	AC016675	AC016675 Homo sapi
27	578.4	40.1	121272	2	AC004956	AC004956 Homo sapi
28	573.6	39.8	186161	2	AC151887	AC151887 Salmiri s
29	567.6	39.4	170630	2	AC016905	AC016905 Homo sapi
30	567.6	39.4	203300	9	AL137073	AL137073 Human DNA
31	562.4	39.0	209421	9	AC090543	AC090543 Homo sapi
32	561.4	38.9	174521	2	AC127468	AC127468 Papio anu
33	559.2	38.8	110459	9	AC117378	AC117378 Homo sapi
34	558.2	38.7	129169	2	AC032038	AC032038 Homo sapi
35	556.8	38.6	48478	2	AC026687	AC026687 Homo sapi
36	556.8	38.6	137737	9	AC010378	AC010378 Homo sapi
37	556.8	38.6	141620	2	AC142444	AC142444 Homo sapi
38	556.8	38.6	149562	2	AC143353	AC143353 Homo sapi
39	556.8	38.6	165736	2	AC144352	AC144352 Homo sapi
40	556.4	38.6	192096	9	AL590133	AL590133 Human DNA
41	555	38.5	91927	9	AC004771	AC004771 Homo sapi
42	554.6	38.5	42665	9	AC011559	AC011559 Homo sapi
43	552	38.3	177166	2	AC123979	AC123979 Papio anu
44	551.8	38.3	211063	2	AC069254	AC069254 Homo sapi
45	550.8	38.2	155949	9	AC091022	AC091022 Homo sapi

## ALIGNMENTS

RESULT 1  
AR051550  
LOCUS AR051550 1442 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 120 from patent US 5830670.  
ACCESSION AR051550  
VERSION AR051550.1 GI:5974914  
KEYWORDS Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1442)  
AUTHORS de la Monte, S. and Wanda, J.R.  
TITLE Neural thread protein gene expression and detection of Alzheimer's disease  
JOURNAL Patent: US 5830670-A 120 03-NOV-1998;  
FEATURES Location/Qualifiers  
source 1..1442  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match	100.0%;	Score 1442;	DB 6;	Length 1442;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1442;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	TTTTTTTTTGGAGATGGAGTTTTCGCTCTTGTTCGCCAGGCTGGAGTGCATGGCGAA	60	
Db	1	TTTTTTTTTGGAGATGGAGTTTTCGCTCTTGTTCGCCAGGCTGGAGTGCATGGCGAA	60	
Qy	61	TCTCAGCTCAGCGCAACCTCCGCCCTCCCGGGTTCAGCGATTTCTCTGCCTCAGCCTCC	120	
Db	61	TCTCAGCTCAGCGCAACCTCCGCCCTCCCGGGTTCAGCGATTTCTCTGCCTCAGCCTCC	120	
Qy	121	CAGTAGCTGGGATACAGGCATGTGCACCCAGCTCGGCTAATTTTGTATTTTTTTTAG	180	
Db	121	CAGTAGCTGGGATACAGGCATGTGCACCCAGCTCGGCTAATTTTGTATTTTTTTTAG	180	
Qy	181	TAGAGATGGAGTTTCTCCATGTGTGTGCTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC	240	
Db	181	TAGAGATGGAGTTTCTCCATGTGTGTGCTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC	240	
Qy	241	CCTCGTCTCGGCTCTCCCAAGTGTAGATACAGGACTGGCCACCATGCCGGCTCTGCC	300	







```

541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGACCACTACACCTGGCTAAATTTTAA 600
601 TTTTATTTTAAATTTTGTGAGACAGATCTCACTCTGACCCAGGCTGAGTGCAAT 660
601 TTTTATTTTAAATTTTGTGAGACAGATCTCACTCTGACCCAGGCTGAGTGCAAT 660
661 GCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCA 720
661 GCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCA 720
721 GCTCTCTGAGTAGCTGGGACTACAGGCGGCCACACGCTAGCTAAATTTTGTATTTT 780
721 GCTCTCTGAGTAGCTGGGACTACAGGCGGCCACACGCTAGCTAAATTTTGTATTTT 780
781 TAGTAGAGATGGGTTTCAACATGTTCCGACAGGTTGATCTTGATCTCTGACCTTGTGATC 840
781 TAGTAGAGATGGGTTTCAACATGTTCCGACAGGTTGATCTTGATCTCTGACCTTGTGATC 840
841 TGCCTGCTCGGCTCCCAAGTGTCTGGGATACAGGCGTGAGCCACACCGCCGCTTAA 900
841 TGCCTGCTCGGCTCCCAAGTGTCTGGGATACAGGCGTGAGCCACACCGCCGCTTAA 900
901 TTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTTACCCAGGCTGAGTGCAAT 960
901 TTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTTACCCAGGCTGAGTGCAAT 960
961 GCGCAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGGATCTCTGCTCA 1020
961 GCGCAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGGATCTCTGCTCA 1020
1021 GCTCTCCAGAGCTGGGATACGGGCACTGCGACCAACACCCCGCTAAATTTTGTATTT 1080
1021 GCTCTCCAGAGCTGGGATACGGGCACTGCGACCAACACCCCGCTAAATTTTGTATTT 1080
1081 TCAATAGAGCGGGTTTCAACATATTTGTCAGGCTGCTCAAACTCTGACCTCAGGT 1140
1081 TCAATAGAGCGGGTTTCAACATATTTGTCAGGCTGCTCAAACTCTGACCTCAGGT 1140
1141 GACCACTGCTCAGCTTCCAAAGTGTGGGATACAGGCGTGAGCCACCTCAACCG 1200
1141 GACCACTGCTCAGCTTCCAAAGTGTGGGATACAGGCGTGAGCCACCTCAACCG 1200
1201 CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGCTTGTGCTATGTGCGCAGGCT 1260
1201 CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGCTTGTGCTATGTGCGCAGGCT 1260
1261 GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACACCCAGCGAGTCA 1320
1261 GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACACCCAGCGAGTCA 1320
1321 CATTTTAAACAGTTACATCTTTATTTAGTATAGTAAAGTAAATACAAATATGT 1380
1321 CATTTTAAACAGTTACATCTTTATTTAGTATAGTAAAGTAAATACAAATATGT 1380
1381 CAAACTGCAATTCAGTAGTACAGAGTTCTTTTATTAATCTTTTAAACAAAGCTTTAGAG 1440
1381 CAAACTGCAATTCAGTAGTACAGAGTTCTTTTATTAATCTTTTAAACAAAGCTTTAGAG 1440
1441 CA 1442
1441 CA 1442

```

```

RESULT 3
AR073235
LOCUS AR073235 1442 bp DNA linear PAT 28-AUG-2000
DEFINITION Sequence 120 from patent US 5948888.
ACCESSION AR073235
VERSION AR073235.1 GI:9999998
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.

```

```

Unclassified.
1 (bases 1 to 1442)
de la Monte, S. and Wands, J. R.
Neural thread protein gene expression and detection of Alzheimer's
disease
Patent: US 5948888-A 120 07-SBP-1999;
Location/Qualifiers
1..1442
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 1442; DB 6; Length 1442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTATTTTGTGAGATGGAGTTTTCGCTCTGTTGCTCCAGGCTGAGTGCAATGGCGAA 60
DB 1 TTTTATTTTGTGAGATGGAGTTTTCGCTCTGTTGCTCCAGGCTGAGTGCAATGGCGAA 60
QY 61 TCTCAGCTCACCGCAACCTCCGCTCCCGGGTTCAAGCGATTTCTCTGCTCAGCTCCC 120
DB 61 TCTCAGCTCACCGCAACCTCCGCTCCCGGGTTCAAGCGATTTCTCTGCTCAGCTCCC 120
QY 121 CAGTAGCTGGGATTAACAGGCAATGTCACACCGCTCGGCTAAATTTTGTATTTTGTAG 180
DB 121 CAGTAGCTGGGATTAACAGGCAATGTCACACCGCTCGGCTAAATTTTGTATTTTGTAG 180
QY 181 TAGAGATGGAGTTTCTCAGTGTGCTCAGGCTGCTGAACTCCGAGCTCAGATGATC 240
DB 181 TAGAGATGGAGTTTCTCAGTGTGCTCAGGCTGCTGAACTCCGAGCTCAGATGATC 240
QY 241 CCTCGCTCTCGGCTCCCAAGAGTGTAGATACAGGCTGGCCACCATGCCGCTCTGCC 300
DB 241 CCTCGCTCTCGGCTCCCAAGAGTGTAGATACAGGCTGGCCACCATGCCGCTCTGCC 300
QY 301 TGGCTAAATTTTGTGTTAGAAACAGGTTTCACTGATGTGCCCAAGCTGGTCTCTCTGAGC 360
DB 301 TGGCTAAATTTTGTGTTAGAAACAGGTTTCACTGATGTGCCCAAGCTGGTCTCTCTGAGC 360
QY 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGCGGATTAACAGGCTGAGCGCTGC 420
DB 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGCGGATTAACAGGCTGAGCGCTGC 420
QY 421 CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTTCCCACTCTTACCCAGGATGAAG 480
DB 421 CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTTCCCACTCTTACCCAGGATGAAG 480
QY 481 TGCAGTGTGTGATCACAAGCTCAGCTCAGCTTCAACTCTCTGAGATCAAGCAATCTCTCTG 540
DB 481 TGCAGTGTGTGATCACAAGCTCAGCTCAGCTTCAACTCTCTGAGATCAAGCAATCTCTCTG 540
QY 541 CCTCAGCTCCCAAGTAGCTGGGACCAAGACATGCACTACACCTGGGCTAAATTTTAA 600
DB 541 CCTCAGCTCCCAAGTAGCTGGGACCAAGACATGCACTACACCTGGGCTAAATTTTAA 600
QY 601 TTTTATTTTAAATTTTGTGAGACAGATCTCACTCTGCTCAGGCTGAGTGCAAT 660
DB 601 TTTTATTTTAAATTTTGTGAGACAGATCTCACTCTGCTCAGGCTGAGTGCAAT 660
QY 661 GCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCA 720
DB 661 GCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCA 720
QY 721 GCTCTCTGAGTAGCTGGGACTACAGGCGGCCACACGCTAGCTAAATTTTGTATTTT 780
DB 721 GCTCTCTGAGTAGCTGGGACTACAGGCGGCCACACGCTAGCTAAATTTTGTATTTT 780
QY 781 TAGTAGAGATGGGTTTCAACATGTTCCGACAGGTTGATCTTGATCTCTGACCTTGTGATC 840
DB 781 TAGTAGAGATGGGTTTCAACATGTTCCGACAGGTTGATCTTGATCTCTGACCTTGTGATC 840
QY 841 TGCCTGCTCGGCTCCCAAGTGTCTGGGATTAACAGGCGTGAGCCACACCGCCGCTTAA 900

```

Db 841 TGCCTGCTCGGCTCCCAAGTCTGGATTACAGCGTGAGCCACACGCGCGCTTA 900  
Qy 901 TTTTAAATTTTGTGTTGAATGGAATCTCACTCTGTTACCCAGGCTGAGTGCAAT 960  
Db 901 TTTTAAATTTTGTGTTGAATGGAATCTCACTCTGTTACCCAGGCTGAGTGCAAT 960  
Qy 961 GGCCTGCTCGGCTCACTGCAACCTCTGCTCCCGGCTCAAGGATCTCTGCTCA 1020  
Db 961 GGCCTGCTCGGCTCACTGCAACCTCTGCTCCCGGCTCAAGGATCTCTGCTCA 1020  
Qy 1021 GCTCTCCAGCAGCTGGGATACGGGCACTGCGGCTCCAGGCTCAAGGATCTCTGCTCA 1080  
Db 1021 GCTCTCCAGCAGCTGGGATACGGGCACTGCGGCTCCAGGCTCAAGGATCTCTGCTCA 1080  
Qy 1081 TCATTAGAGCGGGGTTTCAACATTTGTCAGGCTGGTCTCAAACTCTGAGCTCAGGT 1140  
Db 1081 TCATTAGAGCGGGGTTTCAACATTTGTCAGGCTGGTCTCAAACTCTGAGCTCAGGT 1140  
Qy 1141 GACCCACCTGCTCAGGCTTCCAAAGTCTGGGATTTACAGGCGTGAGCCACCTCAACCCAG 1200  
Db 1141 GACCCACCTGCTCAGGCTTCCAAAGTCTGGGATTTACAGGCGTGAGCCACCTCAACCCAG 1200  
Qy 1201 CCGGCTAAATTTAGATAAAAAATATAGCAATGGGGGCTTGTCTATGTTGCCAGGCT 1260  
Db 1201 CCGGCTAAATTTAGATAAAAAATATAGCAATGGGGGCTTGTCTATGTTGCCAGGCT 1260  
Qy 1261 GGTCTCAAACTTCTGCTTCAAGCAATCTTCCAAATGAGCCACCAACCCAGCCAGTCA 1320  
Db 1261 GGTCTCAAACTTCTGCTTCAAGCAATCTTCCAAATGAGCCACCAACCCAGCCAGTCA 1320  
Qy 1321 CATTTTAAACAGGTACATCTTTATTTAGTATCTAGAGTAACTAGAGTAACTAGAGT 1380  
Db 1321 CATTTTAAACAGGTACATCTTTATTTAGTATCTAGAGTAACTAGAGTAACTAGAGT 1380  
Qy 1381 CAAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATACTTTTAAACAAAGCTTTAGAG 1440  
Db 1381 CAAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATACTTTTAAACAAAGCTTTAGAG 1440  
Qy 1441 CA 1442  
Db 1441 CA 1442

RESULT 4  
AX594290  
LOCUS AX594290 1442 bp DNA linear PAT 13-PBB-2003  
DEFINITION Sequence 1 from Patent WO02074323.  
ACCESSION AX594290  
VERSION AX594290.1 GI:28375449  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
Auerback, P.  
AUTHORS  
TITLE Methods of using neural thread proteins to treat tumors and other  
conditions requiring the removal or destruction of cells  
JOURNAL Patent: WO 02074323-A 1 26-SEP-2002;  
AVERBACK, Paul (CA)  
FEATURES  
source  
1. .1442  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
15. .1142  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD66493.1"  
/db\_xref="GI:28375450"  
/translation="MEFSLLPRLCNGAISAHRLRLPGSSDSPASPSFVAGITGMC  
THARLLYPLVEMELHVQAGLELPTSDPVSASQSAARYRTGHHARLCIANFCGR

NRVSLMCPSPWSPKLSLPIKMDYRRAVPLGLFILFLRHRCPRTLTODEVOMCD  
HSLQPSPEIKHPASASQVAGTDMHYTWLFIPIFENLRQSLNSVQTQGVQRN  
LGSQPLPPGPKLFSCPLSSWDTPRPRPRLANFVFLVEMGFNFARLLILSGCDL  
PASASQSGITGVSHARLIIFNCLFEMESHVTOAGVQWNLGSLQPLPLGLKRFSC  
LSLPSPWDYGHLPHPANFCIFIRGVSVPYLSGWSQTPDLR"  
ORIGIN  
Query Match 100.0%; Score 1442; DB 6; Length 1442;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TTTTATTTTGTGAGATGGAGTTTTCGCTCTTGTGTCGCCAGGCTGAGTGCAATGGCGCAA 60  
Db 1 TTTTATTTTGTGAGATGGAGTTTTCGCTCTTGTGTCGCCAGGCTGAGTGCAATGGCGCAA 60  
Qy 61 TCTCAGCTCAACCGCAACCTCCGCTCCCGGGTTCGAGGATTTCTCTGCTCAGGCTCC 120  
Db 61 TCTCAGCTCAACCGCAACCTCCGCTCCCGGGTTCGAGGATTTCTCTGCTCAGGCTCC 120  
Qy 121 CAGTAGCTGGGATTTACAGGCAATGTCACCCAGCTCGGCTAATTTTGTATTTTTTTAG 180  
Db 121 CAGTAGCTGGGATTTACAGGCAATGTCACCCAGCTCGGCTAATTTTGTATTTTTTTAG 180  
Qy 181 TAGAGATGGAGTTTCTCCATGTTGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240  
Db 181 TAGAGATGGAGTTTCTCCATGTTGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240  
Qy 241 CTTCCGCTCTCGGCTCCCAAGTGTAGATACAGGCTGGCCACCATGCCCCGCTCTGCC 300  
Db 241 CTTCCGCTCTCGGCTCCCAAGTGTAGATACAGGCTGGCCACCATGCCCCGCTCTGCC 300  
Qy 301 TGGCTAAATTTTGTGTTAGAACAGGTTTCACTGATGTGCCAGCTGGTCTCTCTGAGC 360  
Db 301 TGGCTAAATTTTGTGTTAGAACAGGTTTCACTGATGTGCCAGCTGGTCTCTCTGAGC 360  
Qy 361 TCAACGAGTCCACCTGCTCAGGCTCCCAAGTGTGGGATTTACAGGCGTGAGCGCTGC 420  
Db 361 TCAACGAGTCCACCTGCTCAGGCTCCCAAGTGTGGGATTTACAGGCGTGAGCGCTGC 420  
Qy 421 CTGGCTTTTATTTTATTTTAAAGACAGGTTTCACTGATGTGCCAGCTGGTCTCTCTGAGC 480  
Db 421 CTGGCTTTTATTTTATTTTAAAGACAGGTTTCACTGATGTGCCAGCTGGTCTCTCTGAGC 480  
Qy 481 TGCAGTGGTGTGATCACAGCTCACCTGAGGCTTCAACTCTGAGATCAAGCATCTCTCTG 540  
Db 481 TGCAGTGGTGTGATCACAGCTCACCTGAGGCTTCAACTCTGAGATCAAGCATCTCTCTG 540  
Qy 541 CTTCAGGCTCCCAAGTAGCTGGGACCAAGACATGACCATACCTTGGCTAATTTTAA 600  
Db 541 CTTCAGGCTCCCAAGTAGCTGGGACCAAGACATGACCATACCTTGGCTAATTTTAA 600  
Qy 601 TTTTATTTTATTTTATTTTGTGAGACAGGTTCTCAACTCTGTCACCCAGGCTGAGTGCA 660  
Db 601 TTTTATTTTATTTTATTTTGTGAGACAGGTTCTCAACTCTGTCACCCAGGCTGAGTGCA 660  
Qy 661 GGGCAATCTGGGCTCACTGCAACCTCTGCTCCCGGGTTCAGGTTATTTCTCTGCCCA 720  
Db 661 GGGCAATCTGGGCTCACTGCAACCTCTGCTCCCGGGTTCAGGTTATTTCTCTGCCCA 720  
Qy 721 GCTCTCTGAGTAGCTGGGACTACAGGCGCCCAACCGCTAGCTAATTTTGTATTTT 780  
Db 721 GCTCTCTGAGTAGCTGGGACTACAGGCGCCCAACCGCTAGCTAATTTTGTATTTT 780  
Qy 781 TAGTAGAGATGGGGTTTCAACATGTTCCGCAAGGTTGATCTCTGACCTTGTGATC 840  
Db 781 TAGTAGAGATGGGGTTTCAACATGTTCCGCAAGGTTGATCTCTGACCTTGTGATC 840  
Qy 841 TGCCTGCTCGGCTCCCAAGTGTGGATTTACAGGGGTGAGCCACACCGCCGCTTA 900  
Db 841 TGCCTGCTCGGCTCCCAAGTGTGGATTTACAGGGGTGAGCCACACCGCCGCTTA 900  
Qy 901 TTTTAAATTTTGTGTTGAAATGGAATCTCACTCTGTTACCCAGGCTGAGTGCAAT 960

```
Db 901 TTTTAAATTTTGTGTTGTTGAAATGGAATCTCACTCTGTATTACCCAGGCTGGAGTGCAT 960
Qy 961 GGCAAAATCTGGCTCACTCAACACCTCTGCTCCCGGGCTCAAGCGATTCCTCTGTCTCA 1020
Db 961 GGCAAAATCTGGCTCACTCAACACCTCTGCTCCCGGGCTCAAGCGATTCCTCTGTCTCA 1020
Qy 1021 GCCTCCCAAGCAGCTGGGATTAAGGCACTGCGGCACTGCGGCACTGCGGCACTGCGGCACT 1080
Db 1021 GCCTCCCAAGCAGCTGGGATTAAGGCACTGCGGCACTGCGGCACTGCGGCACTGCGGCACT 1080
Qy 1081 TCATTAGAGCGGGGTTTACCATATTTTTCAGGCTGGTCTCAAACTCTGACCTCAGGT 1140
Db 1081 TCATTAGAGCGGGGTTTACCATATTTTTCAGGCTGGTCTCAAACTCTGACCTCAGGT 1140
Qy 1141 GACCCACCTGCTCAGCTTCCAAAGTCTGCGGATTAACAGGCTGAGCCACCTCAACCCAG 1200
Db 1141 GACCCACCTGCTCAGCTTCCAAAGTCTGCGGATTAACAGGCTGAGCCACCTCAACCCAG 1200
Qy 1201 CCGGCTAAATTTAGATTAATAAATATATAGCAATGAGGAGTCTGCTATGTTGCCAGGCT 1260
Db 1201 CCGGCTAAATTTAGATTAATAAATATATAGCAATGAGGAGTCTGCTATGTTGCCAGGCT 1260
Qy 1261 GGTCTCAAACTCTGCGCTTCAATGCTTCCAAAGTCTGCGGATTAACAGGCTGAGCCACCTCA 1320
Db 1261 GGTCTCAAACTCTGCGCTTCAATGCTTCCAAAGTCTGCGGATTAACAGGCTGAGCCACCTCA 1320
Qy 1321 CATTTTAAACAGTTACATCTTTATTTAGTATATCTAGAAATGAGCAATCAATAAATATG 1380
Db 1321 CATTTTAAACAGTTACATCTTTATTTAGTATATCTAGAAATGAGCAATCAATAAATATG 1380
Qy 1381 CAAACCTGCAAAATTCAGTAGTAAACAGAGTTCTTTTATACTTTTAAACAAAGCTTTAGAG 1440
Db 1381 CAAACCTGCAAAATTCAGTAGTAAACAGAGTTCTTTTATACTTTTAAACAAAGCTTTAGAG 1440
Qy 1441 CA 1442
Db 1441 CA 1442

RESULT 5
AX599994 AX599994 1442 bp DNA linear PAT 14-FEB-2003
LOCUS Sequence 1 from Patent W00234915.
DEFINITION AX599994
ACCESSION AX599994
VERSION AX599994.1 GI:28400096
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Fitzpatrick, J., Averback, P., Focht, M.S. and Bibiano, R.
TITLE Preferred segments of neural thread protein and methods of using
JOURNAL the same
Patent: WO 0234915-A 1 02-MAY-2002;
Nymox Pharmaceuticals Corporation (CA)
FEATURES
source
1. .1442
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/notes="NTP DNA sequence"
15. .1142
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD66642.1"
/db_xref="GI:28400097"
/translation="MEFSLLLPRLCNGAISAHNRLPSSDSPASPSFVAGITGMC
THARLLYFPLVEMFELHVGQAGLELPTSDPSVSASQSAARYTGHARLCLIANFCGR
NRVSLMCPSPWPELKOSTCLSLPKWDYRRAAVPGLFLLFRLHRCPTLTQDEVQWCD
HNSLPSTPEIKHPHSPASQVAGTKDMHWYLLIFLFIENFLRQSLNSVTOAGVQRN
LGLQLPQPGKLFSCPSLSSWDYRPRLANFVFLVEMGFTMFARLILSGBCDL
PASASQSGATIGVSHARLIFNFCLEPMESHVSVTQGVQWPNLGLSLQPLPLGRFSC
LSLPSSWDYGHLPHPHPANFCIFIRGVSPYLSGWSQTDLR"
```

## ORIGIN

```
Query Match 100.0%; Score 1442; DB 6; Length 1442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTTAAATTTTGGAGATGGAGTTTTCGCTCTGTGTCAGGCTGGAGTGCATATGGCGCAA 60
Db 1 TTTTAAATTTTGGAGATGGAGTTTTCGCTCTGTGTCAGGCTGGAGTGCATATGGCGCAA 60
Qy 61 TCTCAGCTCACGCGCAACCTCCGCTCCCGGGTTCAGCGATTCCTCTGCTCAGCTCC 120
Db 61 TCTCAGCTCACGCGCAACCTCCGCTCCCGGGTTCAGCGATTCCTCTGCTCAGCTCC 120
Qy 121 CAGTAGCTGGGATTAACAGGCAATGTCACCACTGCGGCTAAATTTTGTATTTTGTAG 180
Db 121 CAGTAGCTGGGATTAACAGGCAATGTCACCACTGCGGCTAAATTTTGTATTTTGTAG 180
Qy 181 TAGAGATGGAGTTTCTCCATGTTGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240
Db 181 TAGAGATGGAGTTTCTCCATGTTGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240
Qy 241 CCTCGCTCTCGGCTCCCAAGTCTAGATACAGGCTGGGCACTGCGGCTCTGCC 300
Db 241 CCTCGCTCTCGGCTCCCAAGTCTAGATACAGGCTGGGCACTGCGGCTCTGCC 300
Qy 301 TGGCTAAATTTTGTGTAGAAACAGGTTTCACTGATGTGCCAAGCTGGTCTCTCGAGC 360
Db 301 TGGCTAAATTTTGTGTAGAAACAGGTTTCACTGATGTGCCAAGCTGGTCTCTCGAGC 360
Qy 361 TCAAGCAGTCCAACCTGCTCAGCTCCCAAGTCTGGGATTAACAGGCTGGGCACTGCGGCT 420
Db 361 TCAAGCAGTCCAACCTGCTCAGCTCCCAAGTCTGGGATTAACAGGCTGGGCACTGCGGCT 420
Qy 421 CTGGGCTTTTATTTTATTTTATTTTAAAGACAGGTTGTCCTTACCCAGGATGAAG 480
Db 421 CTGGGCTTTTATTTTATTTTATTTTAAAGACAGGTTGTCCTTACCCAGGATGAAG 480
Qy 481 TGCAGTGGTGTGATCACAGCTCACTGAGCCTTCAACTCTGAGATCAAGCATCTCTCTG 540
Db 481 TGCAGTGGTGTGATCACAGCTCACTGAGCCTTCAACTCTGAGATCAAGCATCTCTCTG 540
Qy 541 CCTCAGCTCCCAAGTAGCTGGGACCAAGACATGACCACTACACCTGGCTAAATTTT 600
Db 541 CCTCAGCTCCCAAGTAGCTGGGACCAAGACATGACCACTACACCTGGCTAAATTTT 600
Qy 601 TTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 660
Db 601 TTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 660
Qy 661 GGCGCAATCTGCGCTCACCTGCAACCTCTGCTCCCGGGTCAAGTTATTTCTCTGCCCCA 720
Db 661 GGCGCAATCTGCGCTCACCTGCAACCTCTGCTCCCGGGTCAAGTTATTTCTCTGCCCCA 720
Qy 721 GCCTCTCTGAGTAGCTGGGACTACAGCGGCCCAACCACTGCTAGCTAAATTTTGTATTTT 780
Db 721 GCCTCTCTGAGTAGCTGGGACTACAGCGGCCCAACCACTGCTAGCTAAATTTTGTATTTT 780
Qy 781 TAGTAGAGATGGGGTTTACCATGTTGCGCAGGTTGATCTTGTATCTTGAACCTTGTGATC 840
Db 781 TAGTAGAGATGGGGTTTACCATGTTGCGCAGGTTGATCTTGTATCTTGAACCTTGTGATC 840
Qy 841 TGCCTGCTCCGCTCCCAAGTCTGGGATTAACAGGCTGAGCCACCAACCGCGCTT 900
Db 841 TGCCTGCTCCGCTCCCAAGTCTGGGATTAACAGGCTGAGCCACCAACCGCGCTT 900
Qy 901 TTTTAAATTTTGTGTTTGTGTTTGAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAT 960
Db 901 TTTTAAATTTTGTGTTTGTGTTTGAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAT 960
Qy 961 GGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTCCTCTGTCTCA 1020
Db 961 GGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTCCTCTGTCTCA 1020
```

QY 1021 GCCTCCCAAGCAGCTGGGATTAGGGGACCTGCGCACCAACCCCGCTAATTTTGTATTT 1080  
 Db 1021 GCCTCCCAAGCAGCTGGGATTAGGGGACCTGCGCACCAACCCCGCTAATTTTGTATTT 1080  
 QY 1081 TCATTAGAGCGGGGTTTCAACATATTTGTCAAGGCTGGTCTCAAACTCTCGACCTCAGGT 1140  
 Db 1081 TCATTAGAGCGGGGTTTCAACATATTTGTCAAGGCTGGTCTCAAACTCTCGACCTCAGGT 1140  
 QY 1141 GACCCACCTGCTCAGCCTTCCAAAGTGTGGGATTACAGGCGTGAGCCACCTCAACCCAG 1200  
 Db 1141 GACCCACCTGCTCAGCCTTCCAAAGTGTGGGATTACAGGCGTGAGCCACCTCAACCCAG 1200  
 QY 1201 CCGGCTAAATTTAGATAAATAATATAGCAATAGGCGGTCTTGTATGTCGCCAGGCT 1260  
 Db 1201 CCGGCTAAATTTAGATAAATAATATAGCAATAGGCGGTCTTGTATGTCGCCAGGCT 1260  
 QY 1261 GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACACCCAGCAGTCA 1320  
 Db 1261 GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACACCCAGCAGTCA 1320  
 QY 1321 CATTTTAAACAGTTACATCTTTATTTTAGTATACAGAAAGTAAATACATAAACAATGT 1380  
 Db 1321 CATTTTAAACAGTTACATCTTTATTTTAGTATACAGAAAGTAAATACATAAACAATGT 1380  
 QY 1381 CAAACCTGCAATTCAGTAGTAAACAGAGTTCTTTTAACTTTTAAACAAAGCTTTAGAG 1440  
 Db 1381 CAAACCTGCAATTCAGTAGTAAACAGAGTTCTTTTAACTTTTAAACAAAGCTTTAGAG 1440  
 QY 1441 CA 1442  
 Db 1441 CA 1442

RESULT 6  
 AX709024 1442 bp DNA linear PAT 04-APR-2003  
 LOCUS  
 DEFINITION Sequence 48 from Patent WO03008443.  
 ACCESSION AX709024  
 VERSION AX709024.1 GI:29564697  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 Averbach, P.A.  
 Peptides effective in the treatment of tumors and other conditions  
 requiring the removal or destruction of cells  
 Patent: WO 03008443-A 48 30-JAN-2003;  
 Nymox Corporation (CA)  
 Location/Qualifiers  
 1..1442  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 15..1142  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD88177.1"  
 /db\_xref="GI:29564698"  
 CDS  
 /translation="MEFSLILPRLECNCAISAHNRLRPGSSDSPASASPVAGITCMG  
 THARLILYFLVEMEFHLVQAGLEPLTSDSDPSVSNASQSARYRTGSHARLCLANFCGR  
 NRVSLMCPNSWSELKQSTCLSLPKMDYRRAAPVGLFILFRLHRCPTLTQDEVQWCD  
 HSLQSPTEIHPHPASQVAGTKDMHHTWLIIFIFNLRQSLNSVTQAGVQWRN  
 LASLQPLPGFKLSPSLSSWDYRPPRLANFVFLVMGFTMFARLILISGPRDL  
 PASASQAGITGVSHARLIFNFCILPEMESHSVTQAGVQWNLGSLQPLPGKRFKSC  
 LSLPSSWDYGHLPHPANFCIFIRGGSVYLSGWSQTDLR"

ORIGIN

Query Match 100.0%; Score 1442; DB 6; Length 1442;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTATTTTGTGAGATGAGATTTTCGCTCTTGTTGTCGCCAGGCTGAGTGCATGSGGCAA 60  
 Db 1 TTTTATTTTGTGAGATGAGATTTTCGCTCTTGTTGTCGCCAGGCTGAGTGCATGSGGCAA 60  
 QY 61 TCTCAGCTCACCGCAACCTCCGCGGTTCAAGGATTTCTCTGCTCAGGCTCC 120  
 Db 61 TCTCAGCTCACCGCAACCTCCGCGGTTCAAGGATTTCTCTGCTCAGGCTCC 120  
 QY 121 CAGTAGCTGGGATTAACAGCATGTGCAACCCAGCTCGGCTAATTTTGTATTTT 180  
 Db 121 CAGTAGCTGGGATTAACAGCATGTGCAACCCAGCTCGGCTAATTTTGTATTTT 180  
 QY 181 TAGAGATGAGATTTTCTCATGTTGCTCAGGCTGGTCTCGAACTCCGACCTCAGATGTC 240  
 Db 181 TAGAGATGAGATTTTCTCATGTTGCTCAGGCTGGTCTCGAACTCCGACCTCAGATGTC 240  
 QY 241 CCTCCGTCCTCGGCTCCCAAGTGTCTAGATACAGGACTGGCCACCATGCGCGCTCTGCC 300  
 Db 241 CCTCCGTCCTCGGCTCCCAAGTGTCTAGATACAGGACTGGCCACCATGCGCGCTCTGCC 300  
 QY 301 TGGCTAAATTTTGTGTTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTCGAGC 360  
 Db 301 TGGCTAAATTTTGTGTTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTCGAGC 360  
 QY 361 TCAAGCAGTCCACCTCGCTCAGCCTCCCAAGTGTCTGGATTCAGGCGGTGACGCGTGC 420  
 Db 361 TCAAGCAGTCCACCTCGCTCAGCCTCCCAAGTGTCTGGATTCAGGCGGTGACGCGTGC 420  
 QY 421 CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGCTGCCACTCTTACCCAGGATGAAG 480  
 Db 421 CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGCTGCCACTCTTACCCAGGATGAAG 480  
 QY 481 TGCAGTGTGTGATCACTGACGCTTCAACTCTGAGATCAAGCATCTCTCTG 540  
 Db 481 TGCAGTGTGTGATCACTGACGCTTCAACTCTGAGATCAAGCATCTCTCTG 540  
 QY 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGAGCATGCAACCATACCTTGGCTTAAATTTT 600  
 Db 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGAGCATGCAACCATACCTTGGCTTAAATTTT 600  
 QY 601 TTTTATTTTAAATTTTGTGACAGAGTCTCAAATCTGTGTCAACCCAGGCTGAGTGCACT 660  
 Db 601 TTTTATTTTAAATTTTGTGACAGAGTCTCAAATCTGTGTCAACCCAGGCTGAGTGCACT 660  
 QY 661 GCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCCA 720  
 Db 661 GCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCCA 720  
 QY 721 GCCTCTCAGTAGCTGGGACTACAGCGGCCACACAGGCTAGCTAATTTTGTATTTT 780  
 Db 721 GCCTCTCAGTAGCTGGGACTACAGCGGCCACACAGGCTAGCTAATTTTGTATTTT 780  
 QY 781 TAGTAGAGATGGGGTTTCAACATGTTGCGCAGGTGATCTTGATCTCTGGACCTTGTGATC 840  
 Db 781 TAGTAGAGATGGGGTTTCAACATGTTGCGCAGGTGATCTTGATCTCTGGACCTTGTGATC 840  
 QY 841 TGCCTGCTCGGCTTCCCAAGTGTCTGGATTTACAGGCGTGAGCCACACGCGCCGCTT 900  
 Db 841 TGCCTGCTCGGCTTCCCAAGTGTCTGGATTTACAGGCGTGAGCCACACGCGCCGCTT 900  
 QY 901 TTTTATTTTGTGTTGTTGTTGAAATGGATCTCACTCTGTTTACCCAGGCTGAGTGCAAT 960  
 Db 901 TTTTATTTTGTGTTGTTGTTGAAATGGATCTCACTCTGTTTACCCAGGCTGAGTGCAAT 960  
 QY 961 GCGCAAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGGATTTCTCTGCTCA 1020  
 Db 961 GCGCAAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGGATTTCTCTGCTCA 1020  
 QY 1021 GCCTCCCAAGCAGCTGGGATTAACGGGACCTGCGCACCAACCCCGCTAATTTTGTATTT 1080  
 Db 1021 GCCTCCCAAGCAGCTGGGATTAACGGGACCTGCGCACCAACCCCGCTAATTTTGTATTT 1080

QY 1081 TCATTAGAGCGGGGTTTCAACATATTTGTGAGGCTGGTCTCAAACTCCTGACCTCAGGT 1140  
 DB TCAATTAGAGCGGGGTTTCAACATATTTGTGAGGCTGGTCTCAAACTCCTGACCTCAGGT 1140  
 QY 1141 GACCCACCTGCTCAGCCTTCCAAAGTGTGGGATTTACAGGCGTGAGCCACCTCACCAG 1200  
 DB GACCCACCTGCTCAGCCTTCCAAAGTGTGGGATTTACAGGCGTGAGCCACCTCACCAG 1200  
 QY 1201 CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGCTTGTCTATGTGTCGCCAGGCT 1260  
 DB CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGCTTGTCTATGTGTCGCCAGGCT 1260  
 QY 1261 GGTCTCAAACTTCTGGCTCATGCAATCTCTCAAAATGAGCCCAACACCCAGCCAGTCA 1320  
 DB GGTCTCAAACTTCTGGCTCATGCAATCTCTCAAAATGAGCCCAACACCCAGCCAGTCA 1320  
 QY 1321 CATTTTAAACAGTTACATCTTTATTTAGTATAGTATAGTAAGTAATACATAAATATGT 1380  
 DB CATTTTAAACAGTTACATCTTTATTTAGTATAGTATAGTAAGTAATACATAAATATGT 1380  
 QY 1381 CAAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATACTTTTAAACAAAGCTTTAGAG 1440  
 DB CAAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATACTTTTAAACAAAGCTTTAGAG 1440  
 QY 1441 CA 1442  
 DB CA 1442

RESULT 7  
 AX709356  
 LOCUS 1442 bp DNA linear PAT 04-APR-2003  
 DEFINITION Sequence 1 from Patent WO02092115.  
 ACCESSION AX709356  
 VERSION AX709356.1 GI:29564893  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 Averbach, P.A.  
 METHOD OF PREVENTING CELL DEATH USING SEGMENTS OF NEURAL THREAD  
 PROTEINS  
 JOURNAL Patent: WO 02092115-A 1 21-NOV-2002;  
 NYMOX CORPORATION (CA)  
 FEATURES  
 source  
 1..1442  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 15..1142  
 /notes="unassigned protein product"  
 /codon\_start=1  
 /protein\_id="CAD88180.1"  
 /db\_xref="GI:29564894"  
 Translation="MEPSLLPLRLCNGAISAHRLNRLPGSSDPSASPSVAGITGMC  
 THARLLVFLVEMFLVHGOAGLELPTSDPSVASQSARVETGHHARLCLANFCGR  
 NRVSLMCPNSWPELQSTCLSLPKWDYRAAVPGFLIFLRLHRCFLITQDEVQWCD  
 HSLQSPSTLKHPPASASVAGTQDMHHVTLWLFIFLFLRLQSLNSVTOAGVQWRN  
 LGLQPLPFGKLPSCSLSSWYRRPRLANFFVFLVMGFTMFABLLILISGPCDL  
 PASASQAGITGVSHARLIFNFCFLFEMESHVSVOAGVQWPNLQSLQPLPGLKRFSC  
 LSLPSSWDYGLHPHPANFCIFIRGCVSPYLSGWSQTPLDR"

## ORIGIN

Query Match 100.0%; Score 1442; DB 6; Length 1442;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TTTTATTTTTCAGATGGAGTTTTCGCTCTTCTGTCGCCAGGCTGGAGTGCATGGGCA 60  
 DB TTTTATTTTTCAGATGGAGTTTTCGCTCTTCTGTCGCCAGGCTGGAGTGCATGGGCA 60

QY 61 TCTCAGCTCACCGCAACCTCCGCTCCCGGGTTCAAGCGATTTCTCTGCTCAGCCTCCC 120  
 DB TCTCAGCTCACCGCAACCTCCGCTCCCGGGTTCAAGCGATTTCTCTGCTCAGCCTCCC 120  
 QY 121 CAGTAGCTGGGATTTACAGGCATGTGCACCCAGCTCGGCTAAATTTTGTATTTTTTTTAG 180  
 DB CAGTAGCTGGGATTTACAGGCATGTGCACCCAGCTCGGCTAAATTTTGTATTTTTTTTAG 180  
 QY 181 TAGAGATGAGTTTCTCCATGTTGGTTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240  
 DB TAGAGATGAGTTTCTCCATGTTGGTTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240  
 QY 241 CTTCCGCTCTCGGCTCTCCAAAGTGTAGATACAGGACTGGCCCAACATGCGCGCTCTGCC 300  
 DB CTTCCGCTCTCGGCTCTCCAAAGTGTAGATACAGGACTGGCCCAACATGCGCGCTCTGCC 300  
 QY 301 TGGCTAAATTTTGTGGTAGAAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGC 360  
 DB TGGCTAAATTTTGTGGTAGAAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGC 360  
 QY 361 TCAAGCAGTCCACCTGCTCAGCCTCCCAAGTGTGGGATTTACAGGCGTGAGCCGCTGC 420  
 DB TCAAGCAGTCCACCTGCTCAGCCTCCCAAGTGTGGGATTTACAGGCGTGAGCCGCTGC 420  
 QY 421 CTGGCTTTTATTTTATTTTAAAGACAGAGTGTCCCACTCTTACCCAGGATGAAG 480  
 DB CTGGCTTTTATTTTATTTTAAAGACAGAGTGTCCCACTCTTACCCAGGATGAAG 480  
 QY 481 TGCAGTGTGTGATCAGAGTCACTGAGGCTTCAACTCTCTGAGATCAAGCATCTCTCTG 540  
 DB TGCAGTGTGTGATCAGAGTCACTGAGGCTTCAACTCTCTGAGATCAAGCATCTCTCTG 540  
 QY 541 CCTCAGCCTCCCAAGTGTGGGACCAAGAGATGACACCATCTGCTGCTGCTGCTGCTG 600  
 DB CCTCAGCCTCCCAAGTGTGGGACCAAGAGATGACACCATCTGCTGCTGCTGCTGCTG 600  
 QY 601 TTTTATTTTAAATTTTGTGAGACAGAGTCTCAACTCTGTCAACCAGGCTGGAGTGCAGT 660  
 DB TTTTATTTTAAATTTTGTGAGACAGAGTCTCAACTCTGTCAACCAGGCTGGAGTGCAGT 660  
 QY 661 GGGCGAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCCA 720  
 DB GGGCGAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCCA 720  
 QY 721 GCTCTCAGTAGCTGGGACTACAGCGCCCAACAGCTAGCTAAATTTTGTATTTT 780  
 DB GCTCTCAGTAGCTGGGACTACAGCGCCCAACAGCTAGCTAAATTTTGTATTTT 780  
 QY 781 TAGTAGAGATGGGGTTTCAACATGTTGCGCAGGTTGATCTTGATCTCTGGACCTTGTGATC 840  
 DB TAGTAGAGATGGGGTTTCAACATGTTGCGCAGGTTGATCTTGATCTCTGGACCTTGTGATC 840  
 QY 841 TGCCTGCTCGGCTCTCCAAAGTGTGGGATTTACAGGCGTGAGCCACACCGCCCGGCTTA 900  
 DB TGCCTGCTCGGCTCTCCAAAGTGTGGGATTTACAGGCGTGAGCCACACCGCCCGGCTTA 900  
 QY 901 TTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTACCAGGCTGGAGTGCAT 960  
 DB TTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTACCAGGCTGGAGTGCAT 960  
 QY 961 GGGCAAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTTCTCTGTCTCA 1020  
 DB GGGCAAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTTCTCTGTCTCA 1020  
 QY 1021 GCCTCCCAAGCAGCTGGGATTTACGGGCACTTGGCCACCAACCCCGCTAATTTTGTATTT 1080  
 DB GCCTCCCAAGCAGCTGGGATTTACGGGCACTTGGCCACCAACCCCGCTAATTTTGTATTT 1080  
 QY 1081 TCATTAGAGCGGGGTTTTCACATATTTGTGAGGCTGGTCTCAAACTCTGACCTCAGGT 1140  
 DB TCATTAGAGCGGGGTTTTCACATATTTGTGAGGCTGGTCTCAAACTCTGACCTCAGGT 1140  
 QY 1141 GACCCACCTGCTCAGCCTTCCAAAGTGTGGGATTTACAGGCGTGAGCCACCTCACCAG 1200

Db 1141 GACCCACCTGCCTCAGCCTTCCAAAGTGTGGGATTAACAGGCGTGAGCCACCTCACCCAG 1200  
QY 1201 CCGGCTAAATTTAGATAAAAAAATATGTAGCAATGGGGGTCTTGCTATATGTGCCCCAGGCT 1260  
Db 1201 CCGGCTAAATTTAGATAAAAAAATATGTAGCAATGGGGGTCTTGCTATATGTGCCCCAGGCT 1260  
QY 1261 GGTCTCAAACTTCTGGCTTCATGCAATCCTTCCAAATGAGCCACAAACCCAGCCAGTCA 1320  
Db 1261 GGTCTCAAACTTCTGGCTTCATGCAATCCTTCCAAATGAGCCACAAACCCAGCCAGTCA 1320  
QY 1321 CATTTTTTAAACAGTTACATCTTTATTTAGTATATCTAGAAAGTAAATACAATAAACATGT 1380  
Db 1321 CATTTTTTAAACAGTTACATCTTTATTTAGTATATCTAGAAAGTAAATACAATAAACATGT 1380  
QY 1381 CAAACCTGCAAAATTCAGTAGTAAACAGAGTCTTTTTATTAATCTTTTAAACAAAGCTTTAGAG 1440  
Db 1381 CAAACCTGCAAAATTCAGTAGTAAACAGAGTCTTTTTATTAATCTTTTAAACAAAGCTTTAGAG 1440  
QY 1441 CA 1442  
Db 1441 CA 1442

RESULT 8  
AX770807  
LOCUS AX770807 1442 bp DNA linear PAT 02-JUL-2003  
DEFINITION Sequence 125 from Patent WO03008444.  
ACCESSION AX770807  
VERSION AX770807.1 GI:32437975  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Averbach, P.A. and Gemmell, J.  
TITLE Peptides effective in the treatment of tumors and other conditions  
requiring the removal or destruction of cells  
JOURNAL Patent: WO 03008444-A 125 30-JAN-2003;  
NYMOX CORPORATION (CA)  
FEATURES  
Location/Qualifiers  
1..1442  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
15..1142  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAE01151.1"  
/db\_xref="GI:32437976"  
/translation="MEFSLLLPRLECNCAISAHNRLRPGSSDPSASAPVAGITGMC  
THARLILYFLVEMFLVHQAGLELPTSDPSVASQASARYRTGHARLCLIANFCGR  
NRVSLMCPWSPELXQSTCLSLPKCWDYRRAVPGLEFLFLRHRCPRTLQDEVQWCD  
HSLQAPSTPEIKHPSPASQVAGTKDMHYTWLIFIFNPLRQSLNSVTOAGVQVRN  
LSGLAPLPQPKLFPSCPLLSWDYRPRMLANPFLVEMGFTHPARLILISGFCDL  
PASASGAGITVSHHARLIFNFCLEFEMESHVSVTQGVQWPNLGLSLQPLPPLKRPFC  
LSLPSWDYGHLPHPANPFCIFIRGVSPYLSGWSQTPLR"

ORIGIN

Query Match 100.0%; Score 1442; DB 6; Length 1442;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTTATTTTGGAGATGGAGTTTTCGCTCTTGTTGCCAGGCTGGAGTCAATGGCGCAA 60  
Db 1 TTTTATTTTGGAGATGGAGTTTTCGCTCTTGTTGCCAGGCTGGAGTCAATGGCGCAA 60  
QY 61 TCTAGCTCAGCGCAACCTCCGGCTCCCGGTTTCAAGCGATTCTCTGCTCAGCCTCCC 120  
Db 61 TCTAGCTCAGCGCAACCTCCGGCTCCCGGTTTCAAGCGATTCTCTGCTCAGCCTCCC 120  
QY 121 CAGTAGCTGGGATTACAGGATGTGCACCCAGCTCGGCTAATTTTGTATTTTTTTAG 180

Db 121 CAGTAGCTGGGATTACAGGATGTGCACCCAGCTCGGCTAATTTTGTATTTTTTTAG 180  
QY 181 TAGAGATGGAGTTTCTCCATGTTGCTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240  
Db 181 TAGAGATGGAGTTTCTCCATGTTGCTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240  
QY 241 CCTCGTCTCGGCTCTCCAAAGTGTAGATACAGGACTGGGCAACATGCGCGGCTCTGCC 300  
Db 241 CCTCGTCTCGGCTCTCCAAAGTGTAGATACAGGACTGGGCAACATGCGCGGCTCTGCC 300  
QY 301 TGGCTAAATTTTGTGGTAGAACAAGGTTTCACTGATGTGCCAAGCTGGTCTCCTGAGC 360  
Db 301 TGGCTAAATTTTGTGGTAGAACAAGGTTTCACTGATGTGCCAAGCTGGTCTCCTGAGC 360  
QY 361 TCAAGCAGTCCACCTGCTCAGGCTCCCAAAAGTGTGGGATTTACAGGCGTGCAGCGTGC 420  
Db 361 TCAAGCAGTCCACCTGCTCAGGCTCCCAAAAGTGTGGGATTTACAGGCGTGCAGCGTGC 420  
QY 421 CTGGCTTTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG 480  
Db 421 CTGGCTTTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG 480  
QY 481 TGCAGTGTGTGATCACAGCTCACCTGCAGCCTTCAAACTCTGAGATCAAGCATCTCCTG 540  
Db 481 TGCAGTGTGTGATCACAGCTCACCTGCAGCCTTCAAACTCTGAGATCAAGCATCTCCTG 540  
QY 541 CTTACGCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGCTAAATTTT 600  
Db 541 CTTACGCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGCTAAATTTT 600  
QY 601 TTTTATTTTAAATTTTGTGAGACAGATCTCAACTCTGTACCCAGGCTGGAGTGCAGT 660  
Db 601 TTTTATTTTAAATTTTGTGAGACAGATCTCAACTCTGTACCCAGGCTGGAGTGCAGT 660  
QY 661 GSCGCAATCTTGGCTCAGTGCACCACTCTCCCGGTTCAAGTATTTCTCTCCGCCCA 720  
Db 661 GSCGCAATCTTGGCTCAGTGCACCACTCTCCCGGTTCAAGTATTTCTCTCCGCCCA 720  
QY 721 GCCTCTGAGTAGCTGGGACTACAGGCGCCCAACCAAGCTAGTAAATTTTGTATTTT 780  
Db 721 GCCTCTGAGTAGCTGGGACTACAGGCGCCCAACCAAGCTAGTAAATTTTGTATTTT 780  
QY 781 TAGTAGAGATGGGTTTCACTATGTTCCGCGAGTGTGATCTTGATCTCTGACCTTGTGATC 840  
Db 781 TAGTAGAGATGGGTTTCACTATGTTCCGCGAGTGTGATCTTGATCTCTGACCTTGTGATC 840  
QY 841 TGCCTGCTCGGCTCTCCAAAGTGTCTGGGATTTACAGGCGTGAGCCACCAAGCGGCTTA 900  
Db 841 TGCCTGCTCGGCTCTCCAAAGTGTCTGGGATTTACAGGCGTGAGCCACCAAGCGGCTTA 900  
QY 901 TTTTAAATTTTGTGTTGTTGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAT 960  
Db 901 TTTTAAATTTTGTGTTGTTGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAT 960  
QY 961 GSCCAAACTCTCGGCTCAGTGCACCACTCTGCTCCCGGCTCAGCGATTCCTCTCTCA 1020  
Db 961 GSCCAAACTCTCGGCTCAGTGCACCACTCTGCTCCCGGCTCAGCGATTCCTCTCTCA 1020  
QY 1021 GCCTCCCAAGCAGCTGGGATTTACGGGCACCTGSCCAACACACCCCGCTAATTTTGTATTT 1080  
Db 1021 GCCTCCCAAGCAGCTGGGATTTACGGGCACCTGSCCAACACACCCCGCTAATTTTGTATTT 1080  
QY 1081 TCATTAGAGCGGGGTTTCACTATTTTGTAGGCTGTCTCAAACTCTCTGACCTCAGGT 1140  
Db 1081 TCATTAGAGCGGGGTTTCACTATTTTGTAGGCTGTCTCAAACTCTCTGACCTCAGGT 1140  
QY 1141 GACCCACTGCTCAGCTTCCAAAGTGTCTGGGATTTACAGGCTGAGCCACTCACCCAG 1200  
Db 1141 GACCCACTGCTCAGCTTCCAAAGTGTCTGGGATTTACAGGCTGAGCCACTCACCCAG 1200  
QY 1201 CCGGCTAATTTAGATAAAAAAATATGTAGCAATGGGGGTCTTGCTATATGTGCCCCAGGCT 1260



Db 1201 CCGGCTAAATTTAGATAAAAAATATAGTAAATGGGGGCTTGTCTATGTTGCCAGGCT 1260  
 Qy 1261 GGTCTCAAACTTCTGGCTTCATGCAATCTTCAAAATGAGCCACCAACCCAGCAGTCA 1320  
 Db 1261 GGTCTCAAACTTCTGGCTTCATGCAATCTTCAAAATGAGCCACCAACCCAGCAGTCA 1320  
 Qy 1321 CAATTTTAAACAGTTACATCTTTATTTTATTTAGTATATCTAGAAAGTAAATAAACAATGT 1380  
 Db 1321 CAATTTTAAACAGTTACATCTTTATTTTATTTAGTATATCTAGAAAGTAAATAAACAATGT 1380  
 Qy 1381 CAAACCTGCAAAATTCAGTAGTAAACAGAGTTCTTTTATATATCTTTTAAACAAAGCTTTAGAG 1440  
 Db 1381 CAAACCTGCAAAATTCAGTAGTAAACAGAGTTCTTTTATATATCTTTTAAACAAAGCTTTAGAG 1440  
 Qy 1441 CA 1442  
 Db 1441 CA 1442

RESULT 9  
 AX774635  
 LOCUS AX774635 1442 bp DNA linear PAT 09-JUL-2003  
 DEFINITION Sequence 53 from Patent WO02097030.  
 ACCESSION AX774635  
 VERSION AX774635.1 GI:32486166  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1

REFERENCE  
 AUTHORS Averbach, P. A.  
 TITLE Peptides derived from neural thread proteins and their medical use  
 JOURNAL Patent: WO 02097030-A 53 05-DEC-2002;  
 NYMOX CORPORATION (CA)

FEATURES  
 source  
 1..1442  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 15..1142  
 /codon\_start=1  
 /product="AD7C-NTP"  
 /protein\_id="CAE06497.1"  
 /db\_xref="GI:32486167"  
 /translation="MEPSLILPRLCNGAISAHRLRLPGSSDSPASASPVAGITGMC  
 THARLLIFFLVEMFLVHGQAGLELPTSDPSVSASASARYRTHGHARLCLANFCGR  
 NRVSLMCPSPKELQKSLSPKCDYRRAAVPGLFLFLRHRCPPLTDQEVQWCD  
 HSLQPTPEIKHPASASQVAGTDMHHTYTLIFIFNFRQSLNSVQAGVQWNR  
 LGSLOPLPGLFPCPLSLSDYRPRLANPFLVEMGFTMFARLLIISGCDL  
 PASASQAGITGVSHARLIIPNCLPFEMESHVSVOAGVQWPNLQSLPLPGLKRFSC  
 LSLPSSWDYGLHPHPANFICPIRGGVSPYLSGWSQTPDLR"

CDS

ORIGIN

Query Match 100.0%; Score 1442; DB 6; Length 1442;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 TTTTATTTTGGAGTGGAGTTTTCGCTCTGTTGCGCCAGGCTGGAGTGAATGGCGAA 60  
 Db 1 TTTTATTTTGGAGTGGAGTTTTCGCTCTGTTGCGCCAGGCTGGAGTGAATGGCGAA 60  
 Qy 61 TCTCAGCTCACGCAACCTCCGCTCCGGGTTCAAGGATTTCTCTGCTCAGCTCC 120  
 Db 61 TCTCAGCTCACGCAACCTCCGCTCCGGGTTCAAGGATTTCTCTGCTCAGCTCC 120  
 Qy 121 CAGTAGCTGGGATTCAGGATGTCACCCAGCTCGGCTAAATTTTATTTTATTTAG 180  
 Db 121 CAGTAGCTGGGATTCAGGATGTCACCCAGCTCGGCTAAATTTTATTTTATTTAG 180  
 Qy 181 TAGAGATGAGTTTCTCCAGTTGGTTCAGCTGGTCTCGAACTCCGACCTCAGATGATC 240  
 Db 181 TAGAGATGAGTTTCTCCAGTTGGTTCAGCTGGTCTCGAACTCCGACCTCAGATGATC 240

Qy 241 CCTCGCTCTGGGCTCCAAAAGTGTAGATACAGGACTGGGCCAATGCCCCTGCTGCC 300  
 Db 241 CCTCGCTCTGGGCTCCAAAAGTGTAGATACAGGACTGGGCCAATGCCCCTGCTGCC 300  
 Qy 301 TGGCTAAATTTTGGTGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTGAGC 360  
 Db 301 TGGCTAAATTTTGGTGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTGAGC 360  
 Qy 361 TCAAGCAGTCCACCTGCTCAGGCTCCCAAAAGTGTGGGATTTACAGGCGTGCAGCGTGC 420  
 Db 361 TCAAGCAGTCCACCTGCTCAGGCTCCCAAAAGTGTGGGATTTACAGGCGTGCAGCGTGC 420  
 Qy 421 CTGGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTTACCCAGGATGAAG 480  
 Db 421 CTGGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTTACCCAGGATGAAG 480  
 Qy 481 TGCAGTGTGTGATCAGAGTCACTGCGAGCTTCAACTCTCTGAGATCAAGATTCCTCTG 540  
 Db 481 TGCAGTGTGTGATCAGAGTCACTGCGAGCTTCAACTCTCTGAGATCAAGATTCCTCTG 540  
 Qy 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGCACCTACACCTGGCTAAATTTTAA 600  
 Db 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGCACCTACACCTGGCTAAATTTTAA 600  
 Qy 601 TTTTATTTTAAATTTTGGAGACAGAGTCTCAAATCTGTCTCACCAGGCTGGAGTGCAGT 660  
 Db 601 TTTTATTTTAAATTTTGGAGACAGAGTCTCAAATCTGTCTCACCAGGCTGGAGTGCAGT 660  
 Qy 661 GGGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCCA 720  
 Db 661 GGGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCCA 720  
 Qy 721 GCCTCTGAGTAGTGGGACTACAGCGCCCAACAGCTAGCTAAATTTTGTATTTT 780  
 Db 721 GCCTCTGAGTAGTGGGACTACAGCGCCCAACAGCTAGCTAAATTTTGTATTTT 780  
 Qy 781 TAGTAGAGATGGGGTTTCAACATGTTTGGCCAGGTTCATCTTGATCTCTGACCTTGTGATC 840  
 Db 781 TAGTAGAGATGGGGTTTCAACATGTTTGGCCAGGTTCATCTTGATCTCTGACCTTGTGATC 840  
 Qy 841 TGCCTGCTCGGCTCCCAAAAGTGTGGGATTTACAGGCGTGAAGCCACACCGCGCTTAA 900  
 Db 841 TGCCTGCTCGGCTCCCAAAAGTGTGGGATTTACAGGCGTGAAGCCACACCGCGCTTAA 900  
 Qy 901 TTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAT 960  
 Db 901 TTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAT 960  
 Qy 961 GGGCAAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTTCTCTGTCTCA 1020  
 Db 961 GGGCAAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTTCTCTGTCTCA 1020  
 Qy 1021 GCCTCCCAAGCAGCTGGGATTCAGGCTCCGCAACCAACCCCGCTAAATTTTGTATTT 1080  
 Db 1021 GCCTCCCAAGCAGCTGGGATTCAGGCTCCGCAACCAACCCCGCTAAATTTTGTATTT 1080  
 Qy 1081 TCATTAGAGCGGGGTTTCACTATATTTGTGAGGCTGGTCTCAAACTCTGACCTCAGGT 1140  
 Db 1081 TCATTAGAGCGGGGTTTCACTATATTTGTGAGGCTGGTCTCAAACTCTGACCTCAGGT 1140  
 Qy 1141 GACCCACCTGCTCAGGCTTCCAAAGTGTGGGATTTACAGGCGTGAAGCCACCTCACCAG 1200  
 Db 1141 GACCCACCTGCTCAGGCTTCCAAAGTGTGGGATTTACAGGCGTGAAGCCACCTCACCAG 1200  
 Qy 1201 CCGGCTAAATTTAGATAAAAAATATGATAGCAATGGGGGTCTTGCTATGTTGCCAGGCT 1260  
 Db 1201 CCGGCTAAATTTAGATAAAAAATATGATAGCAATGGGGGTCTTGCTATGTTGCCAGGCT 1260  
 Qy 1261 GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACCAACCCAGCAGTCA 1320  
 Db 1261 GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACCAACCCAGCAGTCA 1320

```
QY 1321 CAATTTTAAACAGTTACATCTTTATTTTGTAGTATCTAGAGAGTAATCAATAAACATGT 1380
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1321 CAATTTTAAACAGTTACATCTTTATTTTGTATATCTAGAAAGTAATCAATAAACATGT 1380
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1381 CAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATTAACCTTTTAAACAAAGCTTTAGAG 1440
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1381 CAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATTAACCTTTTAAACAAAGCTTTAGAG 1440
QY 1441 CA 1442
Db ||
QY 1441 CA 1442

RESULT 10
BD070526
LOCUS BD070526
DEFINITION Transgenic animals and cell lines for screening drugs effective for
the treatment or prevention of Alzheimer's disease.
ACCESSION BD070526
VERSION BD070526.1 GI:22616129
KEYWORDS JP 2001513777-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1442)
AUTHORS Monte,S.D. and Wands,J.R.
TITLE Transgenic animals and cell lines for screening drugs effective for
the treatment or prevention of Alzheimer's disease
JOURNAL Patent: JP 2001513777-A 1 04-SEP-2001;
THE GENERAL HOSPITAL CORP
COMMENT OS Unidentified
PN JP 2001513777-A/1
PD 04-SEP-2001
PF 26-FEB-1998 JP 1998537813
PR 26-FEB-1997 US 60/038908
PI SUZANNE DELA MONTE,JACK R WANDS
PC C07H21/02,C07H21/04,C12N5/00,C12N15/00,C12Q1/02,A61K48/00, PC
A61K49/00
CC Strandedness: Double;
CC Topology: Both;
CC Transgenic animals and cell lines for screening drugs CC
CC effective for the
CC treatment or prevention of Alzheimer's disease FH Key
FT Location/Qualifiers
CDS 15..1139.
FEATURES
source
1..1442
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 1442; DB 6; Length 1442;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTTTATTTTGTAGATGAGTTTTCGCTCTGTGTGCCAGGCTCGAGTGCAATGGCGAA 60
Db 1 TTTTATTTTGTAGATGAGTTTTCGCTCTGTGTGCCAGGCTCGAGTGCAATGGCGAA 60
QY 61 TCTCAGCTCACCGCAACCTCCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCCC 120
Db 61 TCTCAGCTCACCGCAACCTCCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCCC 120
QY 121 CAGTAGCTGGGATTACAGGCATGTGCAACCCAGCTCGGCTAATTTTGTATTTTATTTT 180
Db 121 CAGTAGCTGGGATTACAGGCATGTGCAACCCAGCTCGGCTAATTTTGTATTTTATTT 180
QY 181 TAGAGATGAGTTTCTCCATGTTGGTCAGCTGGTCTCGAACTCCCGACCTCAGATGATC 240
Db 181 TAGAGATGAGTTTCTCCATGTTGGTCAGCTGGTCTCGAACTCCCGACCTCAGATGATC 240
QY 241 CCTCGCTCTCGGCTCCCAAAGTGCTAGATACAGGACTGGCCACCATGTCGGCTCTGCC 300
```

```
Db CCTCGCTCTCGGCTCCCAAAGTGCTAGATACAGGACTGGCCACCATGTCGGCTCTGCC 300
QY TGGCTAAATTTTGTGGTAGAAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTCTGAGC 360
Db TGGCTAAATTTTGTGGTAGAAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTCTGAGC 360
QY TCNAGCAGTCCACCTGCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCGTCG 420
Db TCNAGCAGTCCACCTGCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCGTCG 420
QY CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCCCACCTCTTACCCAGGATGAG 480
Db CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCCCACCTCTTACCCAGGATGAG 480
QY TGCAGTGTGTGATCACAGCTCACTGAGCCTTCAACTCTCTGAGATCAAGCATCTCTCTG 540
Db TGCAGTGTGTGATCACAGCTCACTGAGCCTTCAACTCTCTGAGATCAAGCATCTCTCTG 540
QY CCTCAGCCTCCCNAGTAGCTGGGACCAAGACATGCACCATACACCTGGCTAAATTTT 600
Db CCTCAGCCTCCCNAGTAGCTGGGACCAAGACATGCACCATACACCTGGCTAAATTTT 600
QY TTTTATTTTAAATTTTGTAGACAGAGTCTCAACTCTGTCAACCAGGCTGGAGTGCAGT 660
Db TTTTATTTTAAATTTTGTAGACAGAGTCTCAACTCTGTCAACCAGGCTGGAGTGCAGT 660
QY GGCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGTTCAAGTTATTTCTCTGCCCA 720
Db GGCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGTTCAAGTTATTTCTCTGCCCA 720
QY GCCTCTGAGTAGCTGGGACTACAGGCGCCACAGGCTAGCTAATTTTGTATTTT 780
Db GCCTCTGAGTAGCTGGGACTACAGGCGCCACAGGCTAGCTAATTTTGTATTTT 780
QY TAGTAGAGATGGGGTTTCAACATGTTTGGCAGGTGATCTTGATCTCTGACCTCTGTGATC 840
Db TAGTAGAGATGGGGTTTCAACATGTTTGGCAGGTGATCTTGATCTCTGACCTCTGTGATC 840
QY TGCTGCTCGGCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACACGCGCGGCTT 900
Db TGCTGCTCGGCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACACGCGCGGCTT 900
QY TTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTTTACCAGGCTGAGTGCAAT 960
Db TTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTTTACCAGGCTGAGTGCAAT 960
QY GGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGGCTCAAGCGATTCTCTCTCTCA 1020
Db GGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGGCTCAAGCGATTCTCTCTCTCA 1020
QY GCCTCCCAAGCAGCTGGGATTACGGGCACTGCGCACACACCCCGCTAATTTTGTATTT 1080
Db GCCTCCCAAGCAGCTGGGATTACGGGCACTGCGCACACACCCCGCTAATTTTGTATTT 1080
QY TCATTTAGAGCGGGGTTTCAACATATTTGTGAGGCTGGTCTCAAACTCTGACCTCAGGT 1140
Db TCATTTAGAGCGGGGTTTCAACATATTTGTGAGGCTGGTCTCAAACTCTGACCTCAGGT 1140
QY GACCCACCTGCTCAGCCTTCCAAAGTGCTGGGATTACAGGCGTGAGCCACCTCAGCCAG 1200
Db GACCCACCTGCTCAGCCTTCCAAAGTGCTGGGATTACAGGCGTGAGCCACCTCAGCCAG 1200
QY CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGCTTTGCTATGTTGCCAGGCT 1260
Db CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGCTTTGCTATGTTGCCAGGCT 1260
QY GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACACACCCAGCAGTCA 1320
Db GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACACACCCAGCAGTCA 1320
QY CATTTTAAACAGTTTACATCTTTTATTTTGTAGTATCTAGAAAGTAATCAATAAACATGT 1380
Db CATTTTAAACAGTTTACATCTTTTATTTTGTAGTATCTAGAAAGTAATCAATAAACATGT 1380
```



Db 1321 CATTCTTTTAAACAGTTACATCTTTATTTTAGTATATACAGAAAGTAATACAATAAACATGT 1380  
QY 1381 CAAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATAACTTTTAAACAAAGCTTTAGAG 1440  
Db 1381 CAAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATAACTTTTAAACAAAGCTTTAGAG 1440  
QY 1441 CA 1442  
Db 1441 CA 1442  
RESULT 11  
AF010144  
LOCUS Homo sapiens neuronal thread protein AD7c-NTP mRNA linear PRI 02-NOV-2001  
DEFINITION AF010144  
ACCESSION AF010144  
VERSION AF010144.1 GI:3002526  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1442)  
AUTHORS Monte,S.M., Ghanbari,K., Frey,W.H., Beheshti,I., Averbach,P.,  
Hauser,S.L., Ghanbari,H.A. and Wands,J.R.  
TITLE Characterization of the AD7c-NTP cDNA expression in Alzheimer's  
disease and measurement of a 41-kD protein in cerebrospinal fluid  
J. Clin. Invest. 100 (12), 3093-3104 (1997)  
JOURNAL 98064067  
MEDLINE 9399956  
PUBMED 2 (bases 1 to 1442)  
REFERENCE De la Monte,S.M. and Wands,J.R.  
AUTHORS Direct Submission  
TITLE Submitted (24-JUN-1997) MGH Cancer Center, Massachusetts General  
JOURNAL Hospital, 149 13th Street, Room 7308, Charlestown, MA 02129, USA  
FEATURES  
source  
1. .1442  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="neuronal"  
15. .1142  
/note="contains expressed Alu-containing domains; 41 kDa  
phosphoprotein; cell death and neuritic sprouting observed  
in neuronal cells transfected with the AD7c-NTP cDNA"  
/product="neuronal thread protein AD7c-NTP"  
/protein\_id="AAC08737.1"  
/db\_xref="GI:3002527"  
/translation="MEFSLPLRLCNGAISAHRNLRPGSSDPSASAPVAGITGMC  
THARILYPLVMEFSLVHVGQAGLEPLTDDPSVASOSARVYRTHGHRCLIANFCGR  
NRSLMCPMSPELQSTCLSLPKCDVRRRAVGLFILFELRHRPTLTODEVOMCD  
HVSLLQPSPEIKHPHPASASQVAGTKDMHYTHLIFIFELRQLSNVQAGVWRN  
LGSQPLPFGKLFSCPSLLSSVDRRPLRANFVFLVEMGFTMFARLILISGCDL  
PASASAGITVSHHARLIFNFCFLFEMESHVSHTQAGVQWPNLISLQPLPLKRFSC  
LSLPSWDYGHLPHPANFCIPRIGSVSPYLSGWSQTFDLR"  
ORIGIN  
Query Match 100.0%; Score 1442; DB 9; Length 1442;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTTCTTTTTCAGATGGAGTTTCGCTCTGTGTGCCAGGCTGGAGTGCATATGGCGAA 60  
Db 1 TTTTCTTTTTCAGATGGAGTTTCGCTCTGTGTGCCAGGCTGGAGTGCATATGGCGAA 60  
QY 61 TCTCAGCTCAGCGCAACCTCCGCTCCCGGGTTCAAGGATTTCTCTGCTCAGCTCCC 120  
Db 61 TCTCAGCTCAGCGCAACCTCCGCTCCCGGGTTCAAGGATTTCTCTGCTCAGCTCCC 120  
QY 121 CAGTAGCTGGATTACAGCATGTGACCCAGCTCGGCTAATTTTGTATTTTTTTTAG 180  
Db 121 CAGTAGCTGGATTACAGCATGTGACCCAGCTCGGCTAATTTTGTATTTTTTTTAG 180

QY 181 TAGAGATGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC 240  
Db 181 TAGAGATGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC 240  
QY 241 CCTCGCTCTCGGCTCCCAAAAGTGTAGATACAGAGTCGGCCACCAATGCGCGCTGTC 300  
Db 241 CCTCGCTCTCGGCTCCCAAAAGTGTAGATACAGAGTCGGCCACCAATGCGCGCTGTC 300  
QY 301 TGGCTAAATTTTGTGTAGAACACAGGTTTCACTGATGTGCCACAGCTGGTCTCTGAGC 360  
Db 301 TGGCTAAATTTTGTGTAGAACACAGGTTTCACTGATGTGCCACAGCTGGTCTCTGAGC 360  
QY 361 TCACGACGTCACCTGCTCCTCAGCTCCCAAAAGTGTGGGATTAACAGGCGTGCAGCGTGC 420  
Db 361 TCACGACGTCACCTGCTCCTCAGCTCCCAAAAGTGTGGGATTAACAGGCGTGCAGCGTGC 420  
QY 421 CTGGCTTTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG 480  
Db 421 CTGGCTTTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG 480  
QY 481 TGCAGTGGTGTGATCACAGCTCACTGCGAGCTTCAACTCTCAGATCAAGCATCTCTCTG 540  
Db 481 TGCAGTGGTGTGATCACAGCTCACTGCGAGCTTCAACTCTCAGATCAAGCATCTCTCTG 540  
QY 541 CCTCAGCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGCTAAATTTT 600  
Db 541 CCTCAGCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGCTAAATTTT 600  
QY 601 TTTTATTTTAAATTTTGTAGACAGAGTCTCAACTCTGTCACTCAGGCTGGAGTGCAGT 660  
Db 601 TTTTATTTTAAATTTTGTAGACAGAGTCTCAACTCTGTCACTCAGGCTGGAGTGCAGT 660  
QY 661 GSGCGAATCTGGCTCAGCTGCAACCTCTGCTCCGCGGTTCAAGTATTTCTCTGCCCCA 720  
Db 661 GSGCGAATCTGGCTCAGCTGCAACCTCTGCTCCGCGGTTCAAGTATTTCTCTGCCCCA 720  
QY 721 GCCTCTCAGTAGCTGGGACTACAGCGCCCAACCACTAGCTAATTTTGTATTTT 780  
Db 721 GCCTCTCAGTAGCTGGGACTACAGCGCCCAACCACTAGCTAATTTTGTATTTT 780  
QY 781 TAGTAGAGATGGGGTTCCACATGTTGCGCAGGTGATCTTGATCTCTGACCTTGTGATC 840  
Db 781 TAGTAGAGATGGGGTTCCACATGTTGCGCAGGTGATCTTGATCTCTGACCTTGTGATC 840  
QY 841 TGCCTGCTCGGCTCCCAAAAGTGTGGGATTAACAGGCGTGCAGCCACCGCGCTTA 900  
Db 841 TGCCTGCTCGGCTCCCAAAAGTGTGGGATTAACAGGCGTGCAGCCACCGCGCTTA 900  
QY 901 TTTTAAATTTTGTGTTGTTGAAATGGAAATCTCACTCTGTTTACCAGGCTGGAGTGCAT 960  
Db 901 TTTTAAATTTTGTGTTGTTGAAATGGAAATCTCACTCTGTTTACCAGGCTGGAGTGCAT 960  
QY 961 GGCCTAAATCTCGGCTCAGCTGCACTGCAACCTCTGCTCCCGGGCTCAAGGATTTCTCTG 1020  
Db 961 GGCCTAAATCTCGGCTCAGCTGCAACCTCTGCTCCCGGGCTCAAGGATTTCTCTGCTCA 1020  
QY 1021 GCCTCCCAAGCAGCTGGGATTAACGGGACCTGCGCACCAACCGCTAATTTTGTATTT 1080  
Db 1021 GCCTCCCAAGCAGCTGGGATTAACGGGACCTGCGCACCAACCGCTAATTTTGTATTT 1080  
QY 1081 TCATTAGAGGGGGGTTTCCACATATTTGTAGGCTGGTCTCAAACTCTCTGACCTCAGGT 1140  
Db 1081 TCATTAGAGGGGGGTTTCCACATATTTGTAGGCTGGTCTCAAACTCTCTGACCTCAGGT 1140  
QY 1141 GACCCACCTGCTCAGCTTCAAGAGTGTGGGATTAACAGGCGTGCAGCCACCTCAACCCAG 1200  
Db 1141 GACCCACCTGCTCAGCTTCAAGAGTGTGGGATTAACAGGCGTGCAGCCACCTCAACCCAG 1200  
QY 1201 CCGGCTAATTTAGATAAAAAAATATGTAGCAATGGGGGCTCTTGCTATGTGCCCCAGGCT 1260  
Db 1201 CCGGCTAATTTAGATAAAAAAATATGTAGCAATGGGGGCTCTTGCTATGTGCCCCAGGCT 1260

```

QY 1261 GGTCTCAAACTCTGGCTTCATGCAATCTCTCAATAGAGCACACACCCAGCCAGTCA 1320
Db |||||
QY 1261 GGTCTCAAACTCTGGCTTCATGCAATCTCTCAATAGAGCACACACCCAGCCAGTCA 1320
Db |||||
QY 1321 CATTTTAAACAGTTACATCTTTATTTTAGTATACAGAGTAAATACATAAAACATGT 1380
Db |||||
QY 1321 CATTTTAAACAGTTACATCTTTATTTTAGTATACAGAGTAAATACATAAAACATGT 1380
Db |||||
QY 1381 CAAACCTGCAAAATCAGTAGTACAGAGTCTTTTATATACCTTTTAAACAAAGCTTTAGAG 1440
Db |||||
QY 1381 CAAACCTGCAAAATCAGTAGTACAGAGTCTTTTATATACCTTTTAAACAAAGCTTTAGAG 1440
Db |||||
QY 1441 CA 1442
Db ||
QY 1441 CA 1442
Db ||

RESULT 12
HS886K2/c
LOCUS
DEFINITION
Human DNA sequence from clone RP5-886K2 on chromosome 1p35.1-36.12
Contains elongin A (RNA polymerase II elongation factor),
hydroxymethylglutaryl-CoA lyase, GALE (UDP-galactose-4-epimerase)
ESTs, GSS, STS, CpG island, complete sequence.
ACCESSION
AL031295
VERSION
AL031295.1 GI:4376011
KEYWORDS
HTG; elongin A; GALE; HMGCL; hydroxymethylglutaryl-CoA lyase; RNA
polymerase II elongation factor.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 124001)
Direct Submission
Submitted (05-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Mar 7, 1999 this sequence version replaced gi:4056542.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TrEMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chrl
RP5-886K2 is from the library RPCI-5 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2
This sequence is the entire insert of clone RP5-886K2.

```

---

```

FEATURES
            source
Location/Qualifiers
1..124001
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="RZPD:RPCIP704K02886"
/db_xref="taxon:9606"
/chromosome="1"
/map="p35.1-36.12"
/clone="RPS-886K2"
/clone_lib="RPCI-5"
50..150
/note="AluY repeat: matches 195..302 of consensus"
161..324
/note="LiMA5 repeat: matches 6133..6297 of consensus"
327..594
/note="AluJb repeat: matches 1..294 of consensus"
597..863
/note="AluJo repeat: matches 5..298 of consensus"
868..1170
/note="AluYb8 repeat: matches 1..310 of consensus"
1671..1848
/note="MIR repeat: matches 67..261 of consensus"
2288..2429
/note="AluJo/FRAM repeat: matches 157..300 of consensus"
2430..2714
/note="AluSg repeat: matches 1..287 of consensus"
2728..3063
/note="LiM4 repeat: matches 2200..2535 of consensus"
3385..3686
/note="AluSx repeat: matches 1..310 of consensus"
3709..4050
/note="THE1C repeat: matches 1..371 of consensus"
4064..4212
/note="FRAM repeat: matches 1..155 of consensus"
4358..4506
/note="MERSA repeat: matches 15..182 of consensus"
4788..4878
/note="AluJo repeat: matches 54..149 of consensus"
4879..5178
/note="AluSx repeat: matches 1..302 of consensus"
5181..5484
/note="AluSg repeat: matches 1..305 of consensus"
5485..5653
/note="AluJo repeat: matches 149..304 of consensus"
5654..5788
/note="AluSx repeat: matches 1..132 of consensus"
5789..6089
/note="AluSp repeat: matches 1..297 of consensus"
6090..6260
/note="AluSx repeat: matches 132..312 of consensus"
6282..6580
/note="AluY repeat: matches 1..296 of consensus"
6596..7224
/note="Li2 repeat: matches 1995..2748 of consensus"
7225..7293
/note="AluJb repeat: matches 152..220 of consensus"
7294..7753
/note="Li2 repeat: matches 1543..1995 of consensus"
7756..8048
/note="AluSg repeat: matches 1..279 of consensus"
8050..8358
/note="AluY repeat: matches 1..309 of consensus"
8420..8987
/note="Li2 repeat: matches 873..1539 of consensus"
8998..9090
/note="Li2 repeat: matches 1543..1627 of consensus"
9060..9105
/note="Li2 repeat: matches 827..872 of consensus"
9180..9279
/note="MIR repeat: matches 35..139 of consensus"
9489..9671
/note="MERSA repeat: matches 1..189 of consensus"
10530..10594

```



QY 1012 CCTGTCTCAGCTCCCAAGCAGCTGGGATTAAGGCGACCTGGCCACCAACACCCCGCTAATT 1071  
 DB |||||  
 QY 5067 CCTGTCTCAGCTCCCAAGCAGCTGGGATTAAGGCGACCTGGCC--ACACCCCGCTAATT 5010  
 DB |||||  
 QY 1072 TTGTGATTTTCAATTAGAGCGGGGTTTCAACATATTGTTCAGGCTGGTCTCAAACTCCTG 1131  
 DB |||||  
 QY 5009 TTGTGATTTTCAATTAGAGCGGGGTTTCAACATATTGTTCAGGCTGGTCTCAAACTCCTG 4950  
 DB |||||  
 QY 1132 ACCTCAGGTGACCCACCTCCCTCAGCTTCCAGGCTTCAAAAGTGTGGGATTAACAGGCGTGAGCCAC 1191  
 DB |||||  
 QY 4949 ACCTCAGGTGACCCACCTCCCTCAGCTTCCAGGCTTCAAAAGTGTGGGATTAACAGGCGTGAGCCAC 4890  
 DB |||||  
 QY 1192 CTCACCCAGCGGCTTAAATTAGATAAAAAAATATGTAGCAATGGGGGCTTGTCTATGTT 1251  
 DB |||||  
 QY 4889 CTCACCCAGCGGCTTAAATTAGATAAAAAAATATGTAGCAATGGGGGCTTGTCTATGTT 4830  
 DB |||||  
 QY 1252 GCCCAGGCTGGTCTCAAACTTCTGGCTTCATGCAATCCCTTCCAAATGAGCCACCAACACCC 1311  
 DB |||||  
 QY 4829 GCCCAGGCTGGTCTCAAACTTCTGGCTTCATGCAATCCCTTCCAAATGAGCCACCAACACCC 4770  
 DB |||||  
 QY 1312 AGCCAGTCACTTTTAAACAGTTACATCTTTATTTAGTATATCTAGAAAGTAATACAA 1371  
 DB |||||  
 QY 4769 AGCCAGTCACTTTTAAACAGTTACATCTTTATTTAGTATATCTAGAAAGTAATACAA 4710  
 DB |||||  
 QY 1372 TAAACATGTCAAACCTGCAAAATTCAGTAGTAAACAGAGTTCTTTTATACCTTTTAAACAAA 1431  
 DB |||||  
 QY 4709 TAAACATGTCAAACCTGCAAAATTCAGTAGTAAACAGAGTTCTTTTATACCTTTTAAACAAA 4650  
 DB |||||  
 QY 1432 GCTTTAG 1438  
 DB |||||  
 QY 4649 GTATCAG 4643  
 DB |||||

RESULT 13  
 BD070528  
 LOCUS  
 DEFINITION Transgenic animals and cell lines for screening drugs effective for the treatment or prevention of Alzheimer's disease.  
 ACCESSION BD070528  
 VERSION BD070528.1 GI:22616131  
 KEYWORDS JP 2001513777-A/3.  
 SOURCE unidentified  
 ORGANISM unidentified  
 unclassified.  
 REFERENCE 1 (bases 1 to 1418)  
 AUTHORS Monte,S.D. and Wands,J.R.  
 TITLE Transgenic animals and cell lines for screening drugs effective for the treatment or prevention of Alzheimer's disease  
 JOURNAL Patent: JP 2001513777-A 3 04-SEP-2001;  
 THE GENERAL HOSPITAL CORP  
 COMMENT OS Unidentified  
 PN JP 2001513777-A/3  
 PD 04-SEP-2001  
 PF 26-FEB-1998 JP 1998537813  
 PR 26-FEB-1997 US 60/038908  
 PI SUZANNE DELA MONTE,JACK R WANDS  
 PC C07H21/02,C07H21/04,C12N5/00,C12N15/00,C1201/02,A61K48/00, PC A61K49/00  
 CC Strandedness: Both;  
 CC Topology: Both;  
 CC Transgenic animals and cell lines for screening drugs CC  
 CC effective for the treatment or prevention of Alzheimer's disease FH Key  
 CC Location/Qualifiers  
 FT source 1. .1418  
 FT Location/Qualifiers  
 /organism='Unidentified'.  
 1. .1418  
 /organism='unidentified'  
 /mol\_type='genomic DNA'  
 /db\_xref='taxon:32644'  
 ORIGIN

Query Match 84.8%; Score 1223.4; DB 6; Length 1418;  
 Best Local Similarity 96.9%; Pred. No. 0;  
 Matches 1375; Conservative 0; Mismatches 31; Indels 13; Gaps 12;  
 QY 2 TTTTCTTTTTCAGAGTGGAGTTTTCGCTCTCTGTTGCCAGGCTGGAGTGCATATGGCGCAAT 61  
 DB |||||  
 QY 1 TTTTCTTTTTCAGAGTGGAGTTTTCGCTCTCTGTTGCCAGGCTGGAGTGCATATGGCGCAAT 60  
 DB |||||  
 QY 62 CTCAGCTCACCGCAACCTCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGCCTCCCC 121  
 DB |||||  
 QY 61 CTCAGCTCACCGCAACCTCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGCCTCCCC 120  
 DB |||||  
 QY 122 AGTA-GCTGGGATATACAGCATGTGCAACCCAGCTCGGCTAATTTGTTATTTTTTTAG 180  
 DB |||||  
 QY 121 AGTAGGCTGGGATATACAGCATGTGCA-CCAGCTCGGCTAATTTGTTATTTTTTTAG 179  
 DB |||||  
 QY 181 TAGAGATGGAGTTTCTCCATGTTGGTCAAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240  
 DB |||||  
 QY 180 TAGAGATGGAGTTTCTCCATGTTGGTCAAGGCTGGTCTCGAACT-CCGACCTCAGATGATC 238  
 DB |||||  
 QY 241 CTCCTGCTCTCGGCTCCCAAGTGTAGATACAGGACTGGCCACCATGCCCGG-CTCTGC 299  
 DB |||||  
 QY 239 CTCCTGCTCTCGGCTCCCAAGTGTAGATACAGGACTGGCCACCATGCCCGCTCTGC 298  
 DB |||||  
 QY 300 CTGGCTAATTTTGGGTAGAAACAGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAG 359  
 DB |||||  
 QY 299 CTGGCTAATTTTGGGTAGAAACAGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAG 358  
 DB |||||  
 QY 360 CTCAGCAGTCCACCTCGCTCAGCCTCCCAAGTGTCTGGGATTAACAGGCTGAGCCCGTG 419  
 DB |||||  
 QY 359 CTCAGCAGTCCACCTCGCTCAGCCTCCCAAGTGTCTGGGATTAACAGGCTGAGCCCGTG 418  
 DB |||||  
 QY 420 CCTGGCTTTTATTTTATTTTAAAGACACAGGTGTCCACTCTTACCCAGGATGAA 479  
 DB |||||  
 QY 419 CCTGGCTTTTATTTTATTTTAAAGACACAGGTGTCCACTCTTACCCAGGATGAA 478  
 DB |||||  
 QY 480 GTGCACTGGTGTGATCACAGCTCAGTGGAGCTTCAACTCTCTGAGATCAAGCATCTCT 539  
 DB |||||  
 QY 479 GTGCACTGGTGTGATCACAGCTCAGTGGAGCTTCAACTCTCTGAGATCAAGCATCTCT 537  
 DB |||||  
 QY 540 GCCTCAGCCTCC- AAGTAGCTGGGACCAAGACATGCAACCACTACCTGCTAATTTT 598  
 DB |||||  
 QY 538 GCCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGCAACCACTACCTGCTAATTTT 597  
 DB |||||  
 QY 599 TATTTTATTTTATTTTAAAGACACAGGTGTCCACTCTTACCCAGGCTGGAGTGA 658  
 DB |||||  
 QY 598 TATTTTATTTTATTTTAAAGACACAGGTGTCCACTCTTACCCAGGCTGGAGTGA 657  
 DB |||||  
 QY 659 GTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGTTCAAGTTATTTCTCTGCC 718  
 DB |||||  
 QY 658 GTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGTTCAAGTTATTTCTCTGCC 717  
 DB |||||  
 QY 719 CAGCCTCTCTGAGTAGCTGGGACTACAGGCGCCACACGCTAGCTAATTTTGTATT 778  
 DB |||||  
 QY 718 CAGCCTCTCTGAGTAGCTGGGACTACAGGCGCCACCAAGCTAGCTAATTTTGTATT 777  
 DB |||||  
 QY 779 TTTAGTAGAGATGGGG- TTTACCATGTTGCCAGGTTGAT-CTTGATCTCTGACCTTGT 836  
 DB |||||  
 QY 778 TTTAGTAGAGATGGGGTTTCACTGTTGCCAGGTTGATGCTAGTCTCTGACCTTGT 837  
 DB |||||  
 QY 837 GATCTGCTGCTCGGCTCGGCTTCCCAAGTGTGGGATTTACAGG- CGTGAGCCACACGCCCG 895  
 DB |||||  
 QY 838 GATCTGCTGCTCGGCTCGGCTTCCCAAGTGTGGGATTTACAGGACGTGACGCCACCGCCG 897  
 DB |||||  
 QY 896 GCTTATTTTATTTTGTGTTTGAATGGAATCTCACTCTGTTTACCCAGGCTGGAGT 955  
 DB |||||  
 QY 898 GCTTATTTTATTTTGTGTTTGAATGGAATCTCACTCTGTTTACCCAGGCTGGAGT 957  
 DB |||||  
 QY 956 GCAATGGGCAATCTCGGCTCACTGCAACCTCTGCTCCCGGCTCAAGCATTTCTCTG 1015  
 DB |||||  
 QY 958 GCAATGGGCAATCTCGGCTCACTGCAACCTCTGCTCCCGGCTCAAGCATTTCTCTG 1017  
 DB |||||  
 QY 1016 TCTCAGCCTCCCAAGCAGCTGGGATTAACGGGACCTGCGCCACACACCCCGCTAATTTT 1075  
 DB |||||

```
Db 1018 TCTCAGCCTCCCAAGCAGCTGGGATTAACGGGACCTG-CACCACACCCCGCTAAATTTTG 1076
Qy 1076 TATTTTCATTAGAGCGGGGTTTCCACATATTTGTGAGGCTGTCTCAAACTCCTGACCT 1135
Db 1077 TATTTTCATTAGAGCGGGGTTTCCACATATTTGTGAGGCTGTCTCAAACTCCTGACCT 1136
Qy 1136 CAGGTGACCCACCTGCTCAGCCTTCCAAAGTGTGGGATTAACAGGCTGAGCCACCTCA 1195
Db 1137 CAGGTGACCCACCTGCTCAGCCTTCCAAAGTGTGGGATTAACAGGCTGAG--CGCTCA 1194
Qy 1196 CCCAGCGGCTAATTTAGATAAAAAATATGAGCAATGGGGGTTCTGCTATGTTGCC 1255
Db 1195 CCCAGCGGCTAATTTAGATAAAAAATATGAGCAATGGGGGTTCTGCTATGTTGCC 1254
Qy 1256 AGCTGTGCTCAAACTTCTGGGTTCAATGCAATCTTCCAAATGAGCCACACCCAGCC 1315
Db 1255 AGCTGTGCTCAAACTTCTGGGTTCAATGCAATCTTCCAAATGAGCCACACCCAGCC 1314
Qy 1316 AGTCACATTTTTTAAACAGTTACATCTTTATTTAGTATAGTAAAGTAATACATATAA 1375
Db 1315 AGTCACA-TTTTTAAACAGTTACATCTTTATTTAGTATAGTAAAGTATAGATAC 1373
Qy 1376 CATGTCAACCTGCAAAITTCAGTAGTAACAGATCTTT 1414
Db 1374 ATGGCGGAACCTGCAAAITTCAGTAGTAACAGATCTTT 1412

RESULT 14
AR051479
LOCUS
DEFINITION
Sequence 49 from patent US 5830670.
ACCESSION
AR051479
VERSION
AR051479.1 GI:5974843
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 1381)
AUTHORS
de la Monte, S. and Wands, J. R.
TITLE
Neural thread protein gene expression and detection of Alzheimer's
disease
JOURNAL
Patent: US 5830670-A 49 03-Nov-1998;
FEATURES
Location/Qualifiers
source
1..1381
/mol_type="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 74.9%; Score 1080.2; DB 6; Length 1381;
Best Local Similarity 94.4%; Pred. No. 5.2e-309;
Matches 1317; Conservative 0; Mismatches 53; Indels 25; Gaps 18;

Qy 2 TTTTTCATTAGAGGAGTTTCGCTCTGTTGCCAGGCTGAGTGAATGGCGCAAT 61
Db 1 TTTTTCATTAGAGGAGTTTCGCTCTGTTGCCAGGCTGAGTGAATGGCGCAAT 60
Qy 62 CTCAGTCAACCCACCTCCGCTCCGGGTTCAAGCATTTCTCCTGCTCAGCTCC 121
Db 61 CTCAGTCAACCCACCTCCGCTCCGGGTTCAAGCATTTCTCCTGCTCAGCTCC 120
Qy 122 AGTAGCTGGGATTTACAGGCATGTGACCCACGCTCGGCTAAATTTTGTATTTT 181
Db 121 AGTAGCTGGGATTTACAGGCATGTGCA-CCAGCTCGGCTAAATTTTGTATTT 179
Qy 182 AGAGATGGAGTTT-CTCATGTTGGTCAAGCTGTCTGAACTCCCGACCTCAGATGAT 239
Db 180 AGAGATGGAGTTTAACTCCATGTTGGTCAAGCTGTCTGAACTCCCGACCTCAGATGAT 239
Qy 240 CCTCCGCTCTCGGCTCCCAAGTGTCT---AGATACAGACTGGCCACCATGCCGG-CT 295
Db 240 CTCCCGCTCTCGGCTCCCAAGTGTCTGAGATTAACAGGCATGAGCCACCATGCCGGCT 299
```

```
Qy 296 CTGCTGCTGCTAAATTTTGTGTAGAAACAGGGTTTCACTGATG-TGCCCAAGCTGCTCTC 354
Db 300 CTGCTGCTGCTAAATTTTGTGTAGAAACAGGGTTTCACTGATGTTGGCCCAAGCTGCTCTC 359
Qy 355 CTGAGCTCAAGCAGTCCACCTGCTCAGCCTCCAAAGTGTGGGATTAACAGGCTGCGAG 414
Db 360 CTGAGCTCAAGCAGTCCACCTGCTCAGCCTCCAAAGTGTGGGATTAACAGGCTG-CAG 418
Qy 415 CGTGCCTGCGCTCTTTTATTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGG 474
Db 419 CGTGCCTGCGCTCTTTTATTTATTTTATTTTAAAGACACAGGTGTACCACCTCTTACCCAGG 478
Qy 475 ATGAAGTGCAGTGTGTGATCAAGCTCACTCAGCCTTCAACTCCTGAGATCAAGC-AT 533
Db 479 ATGAAGTGCAGTGTGTGATCAAGCTCACTCAGCCTTCAACTCCTGAGATCAAGCAAT 538
Qy 534 CTTCTGCTCAGCTCCCAAGTGTGGGACCAAGACATGACACCTACACCTGGCTA 593
Db 539 CTTCTGCTCAGCTCCCAAGTGTGGGACCAAGACATGACACCTACACCTGG-TA 597
Qy 594 ATTTTATTTTATTTTATTTTAAATTTTGAACAGAGTCTCAACTCTGTCAACCAAGGCTGGA 653
Db 598 ATTTTATTTTATTTTATTTTAAATTTTGAACAGAGTCTC-ACTCTGTCAACCAAGGCTGGA 656
Qy 654 GTGCAGTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGTTCAAGTTATTTCTCC 713
Db 657 GTGCAGTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGTTCAAGTTATTTCTCC 716
Qy 714 TGCCCCAGGCTCTGAGTAGCTGGGACTACAGGCGCCCAACCAAGCTAGCTAATTTTTT 773
Db 717 TGCCCCAGGCTCTCTGAGTAGCTGGGACTACAGGCGCCCAACCAAGCTAGCTAATTTTTT 776
Qy 774 GTATTTTATGATAGATGGGG-TTCAACATGTTTCCAGAGTTGATCTTCACTCTCTGAGACC 832
Db 777 GTATTTTATGATAGATGGGGTTTCAACATGTTTCCAGAGTTGATCTTCACTCTCTTGACC 836
Qy 833 TTGTGATCTGCTGCTCGGCT-CCCAAGTGTGGGATTAACAGGCTGAGCCACACAG 891
Db 837 TTGTGATCTGCTGCTCGGCTACCCAAAGTGTGGGATTAACAG-GTGGTACTCCAC 894
Qy 892 CCGGCTTATTTTATTTTATTTTGTGTTGAATGGAATCTCACTCTGTATCCAGGCTG 951
Db 895 CCGGCTTATTTTATTTTATTTTGTGTTGAATGGAATCTCACTCTGTATCCAGGCTG 954
Qy 952 GAGTGCAATGGCAAACTCTCGCTCACTGCAACCTCTGCTCCCGGCTCAAGCGATTTCT 1011
Db 955 GAGTGCAATGG-CAAACTCTCGCTCACTGCAACCTCTGCTCCCGG-TCAAGCGATTTCT 1012
Qy 1012 CTTGCTCAGCTCCCAAGCAGCTGGGATTAACGGGCACTGCCACACACCCCGCTAAT 1071
Db 1013 CTTGCTCAGCTCCCAAGCAGCTGGGATTAACGGG-ACCTGCACACACCCCGCTAAT 1070
Qy 1072 TTTGTATTTTCAATTAGAGCGGGTTTCAACATATTTGTGAGGCT-GGTCTCAAACTCT 1130
Db 1071 TTTGTATTTTCAATTAGAGCGGG--TTTACCAATATTTGTGAGGCTGGGTTCTCAAACTCT 1128
Qy 1131 GACCTCAGTGCACCACTGCTCAGCCTTCCAAAGTGTGGGATTAACAGGCTGAGCCA 1190
Db 1129 GACCTCAGTGCACCACTGCTCAGCCTTCCAAAGTGTGGGATTAACAGGCTGAGCCA 1188
Qy 1191 CTTCAACCCAGCGCTAATTTAGATAAAAAATATGTAGCAATGGGGGTTTGTGTATGT 1250
Db 1189 CTTCAACCCAGCGCTAATTTTGAATATAAAAAATATGTAGCAATGGGG--TCTGTATGT 1246
Qy 1251 TGCCCCAGGCTGTCTCAAACTTCTGCTTCATGCAATCTTCCAAATGAGCCACACACCC 1310
Db 1247 TGCCCCAGGCTGTCTCAAACTTCTGCTTCATGCAATCTTCCAAATGAGCCACACACCC 1306
Qy 1311 CAGCCAGTCAATTTTTTAAACAGTTACATCTTTTATTTTATGATATACTAGAAAGTATACA 1370
Db 1307 CAGCCAGTCAATTTTTTAAACAGTTACATCTTTTATTTATGATATACTAGAAAGTATACA 1366
Qy 1371 ATAAACATGTCAAAAC 1385
```

```
Db      1367 ATAAACATGTCAAAC 1381
|||||
RESULT 15
LOCUS   AR072619          1381 bp      DNA      linear      PAT 28-AUG-2000
DEFINITION Sequence 49 from patent US 5948634.
ACCESSION AR072619
VERSION   AR072619.1  GI:9999383
KEYWORDS SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS de la Monte,S. and Wands,J.R.
TITLE   Neural thread protein gene expression and detection of alzheimer's
        disease
JOURNAL Patent: US 5948634-A 49 07-SEP-1999;
FEATURES source
        1..1381
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
Query Match      74.9%; Score 1080.2; DB 6; Length 1381;
Best Local Similarity 94.4%; Pred. No. 5.2e-309;
Matches 1317; Conservative 0; Mismatches 53; Indels 25; Gaps 18;
QY      2 TTTTATTTTGTAGATGGAGTTTCGCTCTTGTGTGCCAGGCTGGAGTGCATGGGCGCAAT 61
DB      1 TTTTATTTTGTAGATGGAGTTTCGCTCTTGTGTGCCAGGCTGGAGTGCATGGGCGCAAT 60
QY      62 CTCAGCTCACCGCAACCTCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCCGCC 121
DB      61 CTCAGCTCACCGCAACCTCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCCGCC 120
QY      122 AGTAGCTGGATTACAGGATGTGCACCGCTCGGCTAAATTTTGTATTTTGTAGT 181
DB      121 AGTAGCTGGATTACAGGATGTGCA-CCACGCTCGGCTAAATTTTGTATTTTGTAGT 179
QY      182 AGAGATGGAGTTT--CTCATGTTGGTCAGGCTGTCTCGAACTCCCGACCTCAGATGAT 239
DB      180 AGAGATGGAGTTTAACTCCATGTTGGTCAGGCTGTCTCGAACTCCCGACCTCAGATGAT 239
QY      240 CCTCCGCTCGGCTCCCAAGTGCT---AGATACAGGACTGGCCACCATGCCGG-CT 295
DB      240 CTCCGCTCTCGGCTCGCCCAAGTGCTGAGATTACAGGCAATGAGCCACCATGCCGGCT 299
QY      296 CTGCGCTGGCTAATTTTGTGGTAGAAACAGGGTTTCACTGTATG-TGCCCAAGCTGGTCTC 354
DB      300 CTGCGCTGGCTAATTTTGTGGTAGAAACAGGGTTTCACTGTATGTTGCCCAAGCTGGTCTC 359
QY      355 CTGAGCTCAAGCAGTCCACCTCGCTCAGCCTCCCAAGTGCTGGGATTACAGGCGTGCAG 414
DB      360 CTGAGCTCAAGCAGTCCACCTCGCTCAGCCTCCCAAGTGCTGGGATTACAGGCGT-CA 418
QY      415 CCGTGGCTGGCTTTTATTTATTTTATTTTAAAGACACAGGTTGCCACTCTTACCCAGG 474
DB      419 CCGTGGCTGGCTTTTATTTATTTTATTTTAAAGACACAGGTTGCCACTCTTACCCAGG 478
QY      475 ATGAGATGAGTGGTGTGATACAGCTCACTGCAGCCTTCACTCTGTGATCAAGC-AT 533
DB      479 ATGAGATGAGTGGTGTGATACAGCTCACTGCAGCCTTCACTCTGTGATCAAGCAAT 538
QY      534 CTTCTGCTCAGCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGGCTA 593
DB      539 CTTCTGCTCAGCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGG-TA 597
QY      594 ATTTTATTTTATTTTATTTTATTTTGTAGACAGAGTCTCAACTCTGTCAACCAGGCTGGA 653
DB      598 ATTTTATTTTATTTTATTTTATTTTGTAGACAGAGTCTC-ACTCTGTCAACCAGGCTGGA 656
```

```
QY      654 GTGAGTGGCGCAATCTTTGGCTCACTGCAACCTCTGCCCTCCCGGGTTCAAGTATTCTCC 713
DB      657 GTGAGTGGCGCAATCTTTGGCTCACTGCAACCTCTGCCCTCCCGGGTTCAAGTATTCTCC 716
QY      714 TGCCCCAGCCTCCTGAGTAGCTGGGACTTACAGGCGCCCAACACGCTAGCTAATTTTTTT 773
DB      717 TGCCCCAGCCTCCTGAGTAGCTGGGACTTACAGGCGCCCAACACGCTAGCTAATTTTTTT 776
QY      774 GTATTTTGTAGTAGATGGGG-TTCACCATGTTGCCAGGTTGATCTTTGATCTCTGGACC 832
DB      777 GTATTTTGTAGTAGATGGGGTTTCAACCATGTTGCCAGGTTGATCTTTGATCTCTTGACC 836
QY      833 TTGTGATCTGCTGCTCGGCT-CCCAAAGTGTGGGATTACAGGCTGAGCCACCAAG 891
DB      837 TTGTGATCTGCTGCTCGGCTACCCAAAGTGTGGGATTACAG--GTGCGTGAATCCAC 894
QY      892 CCGGCTTATTTTAAATTTTGTGTTGAAATGGAATCTCACCTCTGTACCCAGGCTG 951
DB      895 GCGGCTTATTTTAAATTTTGTGTTGAAATGGAATCTCACCTCTGTACCCAGGCTG 954
QY      952 GAGTCAATGCGCAAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTCT 1011
DB      955 GAGTCAATGG-CAAAATCTCGGCTACTCGCAACCTCTGCTCCCGGG-TCAAGCGATTCT 1012
QY      1012 CCTGCTCAGCTCCCAAGCAGCTGGGATTACGGGCACCTGCCACCAACCCCGCTAATT 1071
DB      1013 CCTGCTCAGCTCCCAAGCAGCTGGGATTACGGG--ACCTGCAACCAACCCCGCTAATT 1070
QY      1072 TTTGTATTTTCAATTAGAGCGGGGTTTCCACCATATTTGTCAGGCT-GGTCTCAAACTCCT 1130
DB      1071 TTTGTATTTTCAATTAGAGCGGG--TTTACCATATTTGTCAGGCTGGGTCTCAAACTCCT 1128
QY      1131 GACCTCAGGTGAGCCACCTGCTCAGCTTCCAAAGTCTGGGATTACAGGCTGAGCCA 1190
DB      1129 GACCTCAGGTGAGCCACCTGCTCAGCTTCCAAAGTCTGGGATTACAGGCTGAGCCA 1188
QY      1191 CCTCACCCAGCGCGCTAATTTAGATATAAAATATGTAGCAATGGGGGTCTTGCTATGT 1250
DB      1189 CCTCACCCAGCGCGCTAATTTAGATATAAAATATGTAGCAATGGGG--TCTGCTATGT 1246
QY      1251 TGCCAGGCTGGTCTCAAACTTCTGGCTTCATGCAATCCTTCCAAATGAGCCACCAACACC 1310
DB      1247 TGCCAGGCTGGTCTCAAACTTCTGGCTTCAGTCAATCCTTCCAAATGAGCCACCAACACC 1306
QY      1311 CAGCAGTCAATTTTAAACAGTTACATCTTTATTTTAGTATCTAGAAAGTAAATACA 1370
DB      1307 CAGCAGTCAATTTTAAACAGTTACATCTTTATTTTAGTATCTAGAAAGTAAATACA 1366
QY      1371 ATAAACATGTCAAAC 1385
DB      1367 ATAAACATGTCAAAC 1381
```

Search completed: September 15, 2005, 20:33:34

Job time : 6446 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 14:22:52 ; Search time 842 Seconds  
(without alignments)  
10138.084 Million cell updates/sec

Title: US-09-380-203-1  
Perfect score: 1442  
Sequence: 1 TTTTITTTTTTGGATGGAG.....TTAAACAAGCTTTAGACA 1442

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001as:\*  
5: Geneseq2001bs:\*  
6: Geneseq2002as:\*  
7: Geneseq2002bs:\*  
8: Geneseq2003as:\*  
9: Geneseq2003bs:\*  
10: Geneseq2003cs:\*  
11: Geneseq2003ds:\*  
12: Geneseq2004as:\*  
13: Geneseq2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1442	100.0	1442	6	ABN89470
2	1442	100.0	1442	6	ABN89470
3	1442	100.0	1442	8	ABX08221
4	1442	100.0	1442	8	ABX08221
5	1442	100.0	1442	8	ABX23236
6	1442	100.0	1442	8	ABX23236
7	1442	100.0	1442	10	AD7C-NT
8	1442	100.0	1442	10	AD7C-NT
9	1223.4	84.8	1418	2	AD7C-NT
10	1078.6	74.8	1381	2	AD7C-NT
11	1077	74.7	1381	2	AD7C-NT
12	590.4	40.9	33112	10	ACC85730
13	587.8	40.8	65608	6	ABL62910
14	587.8	40.8	65608	6	ABL64414
15	587.8	40.8	65608	6	ABL67658
16	535.2	37.1	79528	6	ABL50814
17	535.2	37.1	79528	12	ADQ19802
18	524.4	36.4	62944	6	ABL68262
19	524.4	36.4	62944	6	ABL66947
20	517.2	35.9	33747	4	AAK73093

c	21	517.2	35.9	33747	4	AAK69279	Human imm
c	22	513.4	35.6	28805	11	ACN45186	Human gen
c	23	513.4	35.6	43545	6	ABK85018	Human cad
c	24	513	35.6	21470	4	ABK42270	Genomic s
c	25	513	35.6	21470	9	ABK60426	Connectiv
c	26	511.6	35.5	5262	4	AAK71768	Human imm
c	27	511.6	35.5	5262	4	AAK71768	Human imm
c	28	511.4	35.5	60729	12	ADQ97864	Human can
c	29	502.8	34.9	75252	11	ACN44450	Human can
c	30	501.4	34.8	3446	13	ADR07063	Human DNA
c	31	496.4	34.4	115756	8	ACD13448	Full leng
c	32	490.4	34.0	99014	6	ABN96931	Gene #342
c	33	490.2	34.0	117143	12	ADJ74882	Marker ge
c	34	490.2	34.0	130001	12	ADJ96282	Human bre
c	35	490	34.0	51837	13	ABD33453	Human can
c	36	489.6	34.0	76201	12	ADJ62814	Human can
c	37	489.6	34.0	24167	5	ABAI6132	Human ner
c	38	489.6	34.0	59999	13	ADR28249	Human low
c	39	482.8	33.5	276276	11	ACN44350	Human gen
c	40	482.8	33.5	60815	11	ACN43882	Human gen
c	41	482.6	33.5	22235	13	ABD33054	Human can
c	42	481.2	33.4	94400	12	ADP08387	Human gly
c	43	480.6	33.3	178024	12	ADQ97721	Human can
c	44	479.6	33.3	44820	6	AAI919703	Reference
c	45	479.2	33.2	51719	6	AAI31365	52Kb gene

## ALIGNMENTS

RESULT 1  
ABN89470  
ID ABN89470 standard; DNA; 1442 BP.  
AC ABN89470;  
XX

DT 02-SEP-2002 (first entry)

DE Neural thread protein (NTP) encoding nucleotide sequence.

KW Neural thread protein; NTP; Harlil peptide; Alzheimer's disease;  
KW Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;  
KW hypoxia; ischaemia; cerebral infarction; gene; ds.

OS Homo sapiens.

FH Key Location/Qualifiers

CDS 15..1142

FT /\*tag= a

FT /product= "Neural thread protein"

XX WO200234915-A2.

XX 02-MAY-2002.

XX 25-OCT-2001; 2001WO-US042813.

XX 27-OCT-2000; 2000US-00697590.

XX (NYMO-) NYMOX PHARM CORP.

XX Fitzpatrick J, Averback P, Focht MSS, Bibiano R;

XX WPI; 2002-507998/54.

XX P-PSDB; ABB81538.

XX New Harlil peptide sequences of the Neural Thread protein, useful in  
XX therapeutic assays, e.g. as targets for developing drugs for treating  
XX Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing  
XX these diseases.

XX Example 1; Fig 1; 53pp; English.



CC The present invention describes a neural thread protein (NTP) peptide  
 CC having an amino acid sequence selected from ABB81511 to ABB81529 and  
 CC their homologues, which are referred collectively as Harilil peptides (1).  
 CC (1) have neuroprotective, nootropic, vasotropic and cerebroprotective  
 CC activities, and can be used in peptide therapy. The Harilil peptide  
 CC sequences can be used as analogues for NTP in therapeutic or diagnostic  
 CC assays by replacing NTP with the peptide in such an assay. The Harilil  
 CC peptides are also useful as a trap material in a diagnostic or  
 CC therapeutic assay. Therefore, the Harilil peptides are useful in binding  
 CC assays, protein and antibody purification, therapeutics or diagnostics.  
 CC In particular, the peptides are also useful for diagnosing Alzheimer's  
 CC disease, Down's syndrome, neuroectodermal tumours, astrocytomas,  
 CC glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides  
 CC are also useful as targets for drug development for the treatment of  
 CC these diseases. The present sequence encodes a neural thread protein  
 CC given in the exemplification of the present invention  
 XX  
 SQ Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 U; 0 Other;

Query Match 100.0%; Score 1442; DB 6; Length 1442;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTCTTTTTCAGATGGAGTTTTCGCTCTGTTGTCAGGCTGAGTGCAATGGCGAA 60  
 DB 1 TTTTCTTTTTCAGATGGAGTTTTCGCTCTGTTGTCAGGCTGAGTGCAATGGCGAA 60

QY 61 TCTCAGCTCACCGCAACCTCCGCTCCGCGGTTCAAGCGATTCTCTGCTCAGCTCC 120  
 DB 61 TCTCAGCTCACCGCAACCTCCGCTCCGCGGTTCAAGCGATTCTCTGCTCAGCTCC 120

QY 121 CAGTAGCTGGGATTACAGGCATGTGCACCCAGCTCGGCTAATTTTGTATTTTGTAG 180  
 DB 121 CAGTAGCTGGGATTACAGGCATGTGCACCCAGCTCGGCTAATTTTGTATTTTGTAG 180

QY 181 TAGAGATGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC 240  
 DB 181 TAGAGATGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC 240

QY 241 CTCCGCTCTCGGCTCCCAAGTGTCTAGATACAGGACTGGCCACCAATGCCGCTCGCC 300  
 DB 241 CTCCGCTCTCGGCTCCCAAGTGTCTAGATACAGGACTGGCCACCAATGCCGCTCGCC 300

QY 301 TGGCTAATTTTGTGTAGAAACAGGGTTTCTAGTGTGCTGAGTGTGCTGCTGAGC 360  
 DB 301 TGGCTAATTTTGTGTAGAAACAGGGTTTCTAGTGTGCTGAGTGTGCTGCTGAGC 360

QY 361 TCAAGCAGTCCACCTCGCTCAGCTCCCAAGTGTGCGATTACAGGCGTGCAGCGTGC 420  
 DB 361 TCAAGCAGTCCACCTCGCTCAGCTCCCAAGTGTGCGATTACAGGCGTGCAGCGTGC 420

QY 421 CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTTGCCACTCTTACCCAGGATGAAG 480  
 DB 421 CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTTGCCACTCTTACCCAGGATGAAG 480

QY 481 TGCAGTGTGTATCAGCTCAGCTCAGCTTCAACTCTCTGAGATCAAGCATCTCTCTG 540  
 DB 481 TGCAGTGTGTATCAGCTCAGCTCAGCTTCAACTCTCTGAGATCAAGCATCTCTCTG 540

QY 541 CCTCAGCTCCCAAGTGTGGACCAAGACATGCACCACTACACCTCGGCTAAATTTTFA 600  
 DB 541 CCTCAGCTCCCAAGTGTGGACCAAGACATGCACCACTACACCTCGGCTAAATTTTFA 600

QY 601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660  
 DB 601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660

QY 661 GGGCGAATCTTGGCTCAGTGCACCTCTGCTCCGCGGTTCAAGTATTCTCTGCGCCA 720  
 DB 661 GGGCGAATCTTGGCTCAGTGCACCTCTGCTCCGCGGTTCAAGTATTCTCTGCGCCA 720

QY 721 GCTCTCTAGTAGCTGGGACTACAGGCGCCACCAAGCTAGTAAATTTTGTATTTT 780

DB 721 GCTCTCTAGTAGCTGGGACTACAGGCGCCACCAAGCTAGTAAATTTTGTATTTT 780  
 QY 781 TAGTAGAGATGGGGTTCCACATGTTTCCAGAGTTGATCTTGTATCTCTGGACCTTGTGATC 840  
 DB 781 TAGTAGAGATGGGGTTCCACATGTTTCCAGAGTTGATCTTGTATCTCTGGACCTTGTGATC 840  
 QY 841 TGCCTGCTCGGCTCCCAAGTGTGGATTACAGGCGTGCAGCCACCAAGCTTGTGATC 900  
 DB 841 TGCCTGCTCGGCTCCCAAGTGTGGATTACAGGCGTGCAGCCACCAAGCTTGTGATC 900  
 QY 901 TTTTAAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 960  
 DB 901 TTTTAAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 960  
 QY 961 GGGCAAAATCTCGGCTCAGTGCACCTCTGCTCCGCGGTTCAAGCGATTCTCTGCTCTCA 1020  
 DB 961 GGGCAAAATCTCGGCTCAGTGCACCTCTGCTCCGCGGTTCAAGCGATTCTCTGCTCTCA 1020  
 QY 1021 GCTCTCCCAAGCAGCTGGGATTACAGGCGCTGCACCAAGCTTGTGTTGTTGTTGTTGTTGTTGTT 1080  
 DB 1021 GCTCTCCCAAGCAGCTGGGATTACAGGCGCTGCACCAAGCTTGTGTTGTTGTTGTTGTTGTTGTT 1080  
 QY 1081 TCAATTAGAGGGGGGTTTCCACATATTTGTCTGAGGCTGGTCTCAAACTCTCTGACCTCAGGT 1140  
 DB 1081 TCAATTAGAGGGGGGTTTCCACATATTTGTCTGAGGCTGGTCTCAAACTCTCTGACCTCAGGT 1140  
 QY 1141 GACCAACCTGCTCAGCTTCCAAAGTGTGGGATTACAGGCGTGCAGCCACCTCTCAGCCAG 1200  
 DB 1141 GACCAACCTGCTCAGCTTCCAAAGTGTGGGATTACAGGCGTGCAGCCACCTCTCAGCCAG 1200  
 QY 1201 CGGCTAATTTTAGATAAAATAATATGATAGCAATGGGGGCTTGTCTATGTTGCCAGGCT 1260  
 DB 1201 CGGCTAATTTTAGATAAAATAATATGATAGCAATGGGGGCTTGTCTATGTTGCCAGGCT 1260  
 QY 1261 GGTCTCAAACTTCTGGCTTTCATGCAATCTTCCAAATGAGCCACCAAGCTTGTGATGTTGCCAGGCT 1320  
 DB 1261 GGTCTCAAACTTCTGGCTTTCATGCAATCTTCCAAATGAGCCACCAAGCTTGTGATGTTGCCAGGCT 1320  
 QY 1321 CATTTTAAACAGTTTATCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1380  
 DB 1321 CATTTTAAACAGTTTATCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1380  
 QY 1381 CAAACCTTCAAAATTCAGTAGTAAACAGTCTTTTATTTTATTTTATTTTATTTTATTTTATTTT 1440  
 DB 1381 CAAACCTTCAAAATTCAGTAGTAAACAGTCTTTTATTTTATTTTATTTTATTTTATTTTATTTT 1440  
 QY 1441 CA 1442  
 DB 1441 CA 1442

RESULT 2

AAD46671  
 ID AAD46671 standard; DNA; 1442 BP.  
 XX  
 AC AAD46671;  
 XX  
 DT 27-JAN-2003 (first entry)  
 XX  
 DB AD7c-NTP gene.

XX Neural thread protein; NTP; hyperplasia; hypertrophy; arteriosclerosis;  
 KW haemorrhoid; gene therapy; tumour; vascular disease; atherosclerosis;  
 KW inflammatory disease; nutritional deficiency disease; genetic disease;  
 KW autoimmune disease; metabolic disease; traumatic disease; intoxication;  
 KW infectious disease; congenital malformation; enzyme deficiency disease;  
 KW amyloid disease; fibrosis disease; storage disease; radiation disease;  
 KW poisoning; environmental disease; endocrine disease; protein therapy;  
 KW degenerative disease; mechanical disease; AD7c-NTP protein; gene; ds.  
 OS Unidentified.  
 XX  
 XX  
 PH Key Location/Qualifiers



CDS 15..1142  
 FT /\*tag= a  
 FT /product= "AD7c-NTP protein"  
 XX PN WC0200274323-A2.  
 XX PD 26-SEP-2002.  
 XX PF 08-MAR-2002; 2002WO-1B001959.  
 XX PR 08-MAR-2001; 2001US-0273957P.  
 XX PA (AVER/) AVERBACK P.  
 XX PI AVerback P;  
 XX DR WPI: 2002-759864/82.  
 XX DR P-PSDB; AAE29142.  
 XX PT Treating a condition in a patient requiring removal or destruction of  
 PT cells, such as a benign or malignant tumor of a tissue or an inflammatory  
 PT disease, comprises administering a neural thread protein (NTP) or a NTP  
 PT gene to a mammal.  
 XX PS Example 2; Fig 1; 70pp; English.  
 XX CC The invention relates to a method for treating a condition in a patient  
 CC requiring removal or destruction of cells. The method involves  
 CC administering to a mammal a neural thread protein (NTP), or administering  
 CC to a tumour or other target cell a NTP gene, where the expression of the  
 CC NTP gene is induced resulting in expression of the NTP protein. The  
 CC method and NTP are useful for treating a condition in a patient requiring  
 CC removal or destruction of cells, such as a benign or malignant tumour of  
 CC a tissue, a hyperplasia, hypertrophy, or overgrowth of a tissue,  
 CC preferably tonsillar hypertrophy or prostatic hyperplasia, a virally,  
 CC bacterially, or parasitically altered tissue, or a malformation of a  
 CC tissue. Other conditions include a cosmetic modification to a tissue,  
 CC such as removal of unwanted facial hair, warts or unwanted fatty tissue,  
 CC a vascular disease, particularly atherosclerosis or arteriosclerosis,  
 CC haemorrhoids, or varicose veins, an inflammatory disease, autoimmune  
 CC disease, metabolic disease, hereditary/genetic disease, traumatic disease  
 CC or physical injury, nutritional deficiency disease, infectious disease,  
 CC congenital malformation, amyloid disease, fibrosis disease, storage  
 CC disease, enzyme deficiency disease, poisoning, intoxication, degenerative  
 CC disease, radiation disease, environmental disease, endocrine disease or  
 CC mechanical disease. The invention is useful in protein therapy and gene  
 CC therapy. The present sequence is AD7c-NTP gene  
 XX SQ Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 1442; DB 6; Length 1442;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TTTTATTTTGGAGTGGAGTTTCTGCTCTTGTGCCAGGCTGGAGTGAATGGCGCAA 60  
 Db 1 TTTTATTTTGGAGTGGAGTTTCTGCTCTTGTGCCAGGCTGGAGTGAATGGCGCAA 60  
 QY 61 TCTCAGCTCAGCGCAACCTCCGGCTCCGGGTTCAAGCGAATCTCTGCTCAGCCTCCC 120  
 Db 61 TCTCAGCTCAGCGCAACCTCCGGCTCCGGGTTCAAGCGAATCTCTGCTCAGCCTCCC 120  
 QY 121 CAGTAGCTGGGATTACAGCAGTGTGCACCCAGCTCGGCTAAATTTTGTATTTTGTAG 180  
 Db 121 CAGTAGCTGGGATTACAGCAGTGTGCACCCAGCTCGGCTAAATTTTGTATTTTGTAG 180  
 QY 181 TAGAGATGGAGTTTCTCCATTTGTGTGAGGCTGGTCTCGAACTCCCGAATCAGATGATC 240  
 Db 181 TAGAGATGGAGTTTCTCCATTTGTGTGAGGCTGGTCTCGAACTCCCGAATCAGATGATC 240  
 QY 241 CCTCGCTCTCGGCTCCCAAGTGTGTAGATACAGGACTGGCCACCAATGCCGGCTCTGCC 300  
 Db 241 CCTCGCTCTCGGCTCCCAAGTGTGTAGATACAGGACTGGCCACCAATGCCGGCTCTGCC 300

QY 301 TGGCTAAATTTTGTGTGAGAAACAGGGTTTCACTGATGTGCCAAGCTGTGTCTCTGAGC 360  
 Db 301 TGGCTAAATTTTGTGTGAGAAACAGGGTTTCACTGATGTGCCAAGCTGTGTCTCTGAGC 360  
 QY 361 TCAAGCAGTCCACCTGCGCTCAGCCTCCCAAAGTGTGGGATTAACAGGCTGACGCGTGC 420  
 Db 361 TCAAGCAGTCCACCTGCGCTCAGCCTCCCAAAGTGTGGGATTAACAGGCTGACGCGTGC 420  
 QY 421 CTGGGCTTTTAAATTTTAAAGACAGAGTGTCCCACTCTTACCCAGGATGAAG 480  
 Db 421 CTGGGCTTTTAAATTTTAAAGACAGAGTGTCCCACTCTTACCCAGGATGAAG 480  
 QY 481 TGCAGTGTGTGATCACAGCTCACTGCGAGCCTTCAACTCTGAGATCAAGCATCTCTCTG 540  
 Db 481 TGCAGTGTGTGATCACAGCTCACTGCGAGCCTTCAACTCTGAGATCAAGCATCTCTCTG 540  
 QY 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGACCACTACACCTGGCTAAATTTT 600  
 Db 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGACCACTACACCTGGCTAAATTTT 600  
 QY 601 TTTTATTTTAAATTTTAAAGACAGAGTCTCAACTCTGTCTACCCAGGCTGGAGTGCAGT 660  
 Db 601 TTTTATTTTAAATTTTAAAGACAGAGTCTCAACTCTGTCTACCCAGGCTGGAGTGCAGT 660  
 QY 661 GGGCGCAATCTTGGCTCACTGCAACCTTGCCTCCCGGGTTCAAGTTTATTTCTCTGCCCA 720  
 Db 661 GGGCGCAATCTTGGCTCACTGCAACCTTGCCTCCCGGGTTCAAGTTTATTTCTCTGCCCA 720  
 QY 721 GCCTCTCAGTAGTGGGACTACAGGCGCCCAACACGCTAGCTAAATTTTGTATTTT 780  
 Db 721 GCCTCTCAGTAGTGGGACTACAGGCGCCCAACACGCTAGCTAAATTTTGTATTTT 780  
 QY 781 TAGTAGAGATGGGGTTTCAACATGTTTGCAGAGTGTGATCTTGATCTCTGACCTTGTGATC 840  
 Db 781 TAGTAGAGATGGGGTTTCAACATGTTTGCAGAGTGTGATCTTGATCTCTGACCTTGTGATC 840  
 QY 841 TGCTTGCCTCGGCTCCCAAAGTGTGGGATTAACAGGCTGAGCCACACGCGCGGCTTA 900  
 Db 841 TGCTTGCCTCGGCTCCCAAAGTGTGGGATTAACAGGCTGAGCCACACGCGCGGCTTA 900  
 QY 901 TTTTAAATTTTGTGTGAAATGGAATCTCACTGTGTACCAGGCTGGAGTGAAT 960  
 Db 901 TTTTAAATTTTGTGTGAAATGGAATCTCACTGTGTACCAGGCTGGAGTGAAT 960  
 QY 961 GGGCAAAATCTCGGCTCACTGCAACCTTGCCTCCCGGGCTCAAGCGATTTCTCTGTCTCA 1020  
 Db 961 GGGCAAAATCTCGGCTCACTGCAACCTTGCCTCCCGGGCTCAAGCGATTTCTCTGTCTCA 1020  
 QY 1021 GCCTCCCAAGCAGCTGGGATTAACGGGCACTGCGCCACACACCCCGCTAAATTTTGTATTT 1080  
 Db 1021 GCCTCCCAAGCAGCTGGGATTAACGGGCACTGCGCCACACACCCCGCTAAATTTTGTATTT 1080  
 QY 1081 TCATTAGAGGCGGGTTTCAACATATTTGTGAGGCTGGTCTCAAACTCTCTGACCTCAGGT 1140  
 Db 1081 TCATTAGAGGCGGGTTTCAACATATTTGTGAGGCTGGTCTCAAACTCTCTGACCTCAGGT 1140  
 QY 1141 GACCCACCTGCTCAGCCTTCCAAAGTGTGGGATTAACAGGCTGAGCCACCTCACCAG 1200  
 Db 1141 GACCCACCTGCTCAGCCTTCCAAAGTGTGGGATTAACAGGCTGAGCCACCTCACCAG 1200  
 QY 1201 CCGGCTAAATTTAGATAAAAAAATATGTAGCAATGGGGGTCTTGTATGTGCCAGGCT 1260  
 Db 1201 CCGGCTAAATTTAGATAAAAAAATATGTAGCAATGGGGGTCTTGTATGTGCCAGGCT 1260  
 QY 1261 GGTCTCAAACTTCTGGCTTTCATGCAATCTTCCAAATGAGCCACCAACCCAGGCTCA 1320  
 Db 1261 GGTCTCAAACTTCTGGCTTTCATGCAATCTTCCAAATGAGCCACCAACCCAGGCTCA 1320  
 QY 1321 CATTTTTTAAACAGTTACATCTTTATTTTATTTAGTATATACTAGAAAGTAAATACAAATGT 1380  
 Db 1321 CATTTTTTAAACAGTTACATCTTTATTTTATTTAGTATATACTAGAAAGTAAATACAAATGT 1380

QY 1381 CAAACCTGCAATTCAGTAGTAACAGAGTTCTTTATTAACCTTTTAAACAAAGCTTTAGAG 1440  
 Db |||||||  
 QY 1441 CA 1442  
 Db ||  
 1441 CA 1442

RESULT 3

ABX08221  
 ID ABX08221 standard; DNA; 1442 BP.

AC ABX08221;

XX 20-JAN-2003 (first entry)

XX DNA encoding human neural thread protein AD7C-NTP.

XX Neural thread protein; NTP-peptide; AD7C-NTP; surgical excision;  
 KW transplantation; grafting; chemotherapy; immunotherapy; vaccination;  
 KW ablation; cryotherapy; laser therapy; phototherapy; gene therapy;  
 KW radiation; tumour; hyperplasia; hypertrophy; overgrowth of tissue;  
 KW malformation of tissue; tonsillar hypertrophy; prostatic hyperplasia;  
 KW cosmetic modification; vascular disease; atherosclerosis;  
 KW arteriosclerosis; haemorrhoid; varicose vein; inflammatory disease;  
 KW autoimmune disease; metabolic disease; traumatic disease;  
 KW physical injury; nutritional deficiency disease; infectious disease;  
 KW amyloid disease; fibrosis disease; storage disease;  
 KW congenital malformation; enzyme deficiency disease; poisoning;  
 KW intoxication; environmental disease; radiation disease;  
 KW endocrine disease; degenerative disease; mechanical disease; gene; ds.

XX Homo sapiens.

OS WO200297030-A2.

XX 05-DEC-2002.

XX 24-MAY-2002; 2002WO-CA000759.

XX 25-MAY-2001; 2001US-0293156P.

XX (NYMO-) NYMOX CORP.

XX Averbach PA;

XX WPI; 2003-041406/03.

XX Novel peptides similar in amino acid sequence to neural thread proteins  
 PT (NTP), useful for treating unwanted cellular proliferations such as  
 PT malignant tumors and prostatic hyperplasia.

XX Disclosure; Fig 1; 78pp; English.

XX The invention describes an NTP-peptide (I) comprising at least one amino  
 CC acid sequence corresponding to part of the amino acid sequence of a  
 CC neural thread protein, AD7C-NTP. The invention provides a method of  
 CC treating a condition requiring removal or destruction of cells of a  
 CC mammal comprising administering to a mammal, a therapeutic amount of (I).  
 CC The treatment is administered to the mammal before, during or after  
 CC surgical excision, transplantation, grafting, chemotherapy,  
 CC immunotherapy, vaccination, thermal or electrical ablation, cryotherapy,  
 CC laser therapy, phototherapy, gene therapy and/or radiation. The method is  
 CC useful for treatment of benign or malignant tumour; hyperplasia,  
 CC hypertrophy or overgrowth of tissue; virally, bacterially or  
 CC parasitically altered tissue; malformation of tissue selected from lung,  
 CC breast, stomach, pancreas, prostate, bladder, bone, ovary, skin, kidney,  
 CC sinus, colon, intestine, rectum, esophagus, heart, spleen, salivary  
 CC gland, blood, brain and its coverings, spinal cord, muscle, connective  
 CC tissue, adrenal, parathyroid, thyroid, uterus, testis, pituitary,  
 CC reproductive organs, liver, hair, gall bladder, eye, ear, nose, throat,  
 CC tonsils, mouth and lymph nodes and lymphoid system; tonsillary

CC hypertrophy; prostatic hyperplasia; cosmetic modification to a tissue;  
 CC vascular disease (atherosclerosis or arteriosclerosis); haemorrhoids;  
 CC varicose veins; inflammatory disease; autoimmune disease; metabolic  
 CC disease; hereditary/genetic disease; traumatic disease; physical injury;  
 CC nutritional deficiency disease; infectious disease; amyloid disease;  
 CC fibrosis disease; storage disease; congenital malformation; enzyme  
 CC deficiency disease; poisoning; intoxication; environmental disease;  
 CC radiation disease; endocrine disease; degenerative disease and mechanical  
 CC disease. This sequence encodes the human neural thread protein AD7C-NTP

XX Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 U; 0 Other;

QY Query Match 100.0%; Score 1442; DB 8; Length 1442;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTT TTTT TTTT GAGATGGAGTTTTCGCTCTTTGTTGCCAGGCTGGAGTGCATGGCGCAA 60

Db 1 TTTT TTTT TTTT GAGATGGAGTTTTCGCTCTTTGTTGCCAGGCTGGAGTGCATGGCGCAA 60

QY 61 TCTAGCTCAGCGCAACCTCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCC 120

Db 61 TCTAGCTCAGCGCAACCTCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCC 120

QY 121 CAGTAGCTGGGATTACAGGCGATGTCACCGCTCGGCTAATTTTGTATTTTTTTAG 180

Db 121 CAGTAGCTGGGATTACAGGCGATGTCACCGCTCGGCTAATTTTGTATTTTTTTAG 180

QY 181 TAGAGATGGAGTTTCTCCATGTTGCTCAGGCTGGTCTCGAATCTCCGACCTCAGATGATC 240

Db 181 TAGAGATGGAGTTTCTCCATGTTGCTCAGGCTGGTCTCGAATCTCCGACCTCAGATGATC 240

QY 241 CCTCGTCTCGGCTCCCAAAAGTGTAGATACAGACTGGCGCACATGCCGGCTCTGCC 300

Db 241 CCTCGTCTCGGCTCCCAAAAGTGTAGATACAGACTGGCGCACATGCCGGCTCTGCC 300

QY 301 TGGCTAAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCAGCTGCTCTCTGAGC 360

Db 301 TGGCTAAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCAGCTGCTCTCTGAGC 360

QY 361 TCAAGCAGTCCACCTGCTCAGCCTCCCAAAAGTGTGGGATTTACAGGCGTGCAGCGTGC 420

Db 361 TCAAGCAGTCCACCTGCTCAGCCTCCCAAAAGTGTGGGATTTACAGGCGTGCAGCGTGC 420

QY 421 CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCCCACTTTACCCAGGATGAAG 480

Db 421 CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCCCACTTTACCCAGGATGAAG 480

QY 481 TGCAGTGTGTGATCACAGCTCACCTGACGCTTCAACTCTCTGAGATCAAGCATCTCTCTG 540

Db 481 TGCAGTGTGTGATCACAGCTCACCTGACGCTTCAACTCTCTGAGATCAAGCATCTCTCTG 540

QY 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGCTAAATTTTFA 600

Db 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGCTAAATTTTFA 600

QY 601 TTTTATTTTAAATTTTGTAGACAGAGTCTCAACTCTGTCAACAGGCTGGAGTGCAGT 660

Db 601 TTTTATTTTAAATTTTGTAGACAGAGTCTCAACTCTGTCAACAGGCTGGAGTGCAGT 660

QY 661 GCGCAATCTTGGCTCACCTGCAACCTCTGCTCCCGGGTTCAAGTTATTCTCTGCCCA 720

Db 661 GCGCAATCTTGGCTCACCTGCAACCTCTGCTCCCGGGTTCAAGTTATTCTCTGCCCA 720

QY 721 GCTCTCTCAGTAGCTGGGACTACAGCGGCCCAACCACTAGCTAATTTTGTATTTT 780

Db 721 GCTCTCTCAGTAGCTGGGACTACAGCGGCCCAACCACTAGCTAATTTTGTATTTT 780

QY 781 TAGTAGAGATGGGGTTTCAACATGTTTCGCGAGGTGATCTTGATCTCTGACCTTGTATC 840

Db 781 TAGTAGAGATGGGGTTTCAACATGTTTCGCGAGGTGATCTTGATCTCTGACCTTGTATC 840

QY 841 TGCCTGCTCGGCTCCCAAAAGTGTGGGATTTACAGGCGTGTAGCCCAACCGCCGCTTFA 900

841 TGCTGCTCGGCTCCCAAGTGTGGGATTACAGCGTGAGCCACCGCCGCTTA 900  
 901 TTTTAAATTTTGTGTTGAATGGAATCTCACTCTGTGTACCGAGGCTGGAGTCAAT 960  
 901 TTTTAAATTTTGTGTTGAATGGAATCTCACTCTGTGTACCGAGGCTGGAGTCAAT 960  
 961 GGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTCTCTGTCTCA 1020  
 961 GGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTCTCTGTCTCA 1020  
 1021 GCTTCCCAAGCAGCTGGGATTACGGGCACTGTCACACACACCCCGCTAAATTTTGTATTT 1080  
 1021 GCTTCCCAAGCAGCTGGGATTACGGGCACTGTCACACACACCCCGCTAAATTTTGTATTT 1080  
 1081 TCATTAGAGCGGGGTTTCAACCATATTTGTGAGGCTGTCTCAAACTCTGACCTCAGGT 1140  
 1081 TCATTAGAGCGGGGTTTCAACCATATTTGTGAGGCTGTCTCAAACTCTGACCTCAGGT 1140  
 1141 GACCCACCTGCTCAGGCTTCCAAAGTGTGGGATTACAGGCGTGAGCCACCTCACCCAG 1200  
 1141 GACCCACCTGCTCAGGCTTCCAAAGTGTGGGATTACAGGCGTGAGCCACCTCACCCAG 1200  
 1201 CCGGCTAAATTAGATAAAAAATATGTAGCAATGGGGGCTTGTCTATGTGTCAGGCT 1260  
 1201 CCGGCTAAATTAGATAAAAAATATGTAGCAATGGGGGCTTGTCTATGTGTCAGGCT 1260  
 1261 GGTCTCAAACTCTGCTCAGCAATCTTCCAAATGAGCCACACACCCGAGTCA 1320  
 1261 GGTCTCAAACTCTGCTCAGCAATCTTCCAAATGAGCCACACACCCGAGTCA 1320  
 1321 CATTTTTTAAACAGTTACATCTTTATTTAGTATACAGAAAGTAAATACATAAATCT 1380  
 1321 CATTTTTTAAACAGTTACATCTTTATTTAGTATACAGAAAGTAAATACATAAATCT 1380  
 1381 CAAACCTGCAAAATCAGTAGTAACAGAGTCTTTTATTAATTTTAAACAAAGCTTTAGAG 1440  
 1381 CAAACCTGCAAAATCAGTAGTAACAGAGTCTTTTATTAATTTTAAACAAAGCTTTAGAG 1440  
 1441 CA 1442  
 1441 CA 1442

RESULT 4  
 AB223236  
 ID AB223236 standard; DNA; 1442 BP.  
 XX AC AB223236;  
 XX DT 24-MAR-2003 (first entry)  
 XX DE Nucleotide sequence of human neuronal thread protein AD7c-NTP.  
 XX KW Human; neuronal thread protein; AD7c-NTP; Alzheimer's disease; histone;  
 XX KW neurodegeneration; in vivo gene expression; amphipathic compound;  
 XX KW gene therapy; gene; ss.  
 XX OS Homo sapiens.  
 XX FH Key  
 XX CDS 15..1142  
 XX FT /\*tag= a  
 XX FT /product= "neuronal thread protein AD7c-NTP"  
 XX FN WO200299036-A2.  
 XX PD 12-DEC-2002.  
 XX PP 28-MAY-2002; 2002WO-US016429.  
 XX PR 01-JUN-2001; 2001US-00872969.

(RHOD-) RHODE ISLAND HOSPITAL.  
 Wands JR, De La Monte SM;  
 WPI; 2003-140605/13.  
 P-PSDB; ABB99774.  
 Inducing prolonged in vivo gene expression in mammal by contacting neuronal tissue with composition comprising Alzheimer's disease-associated neural thread protein 7c antisense nucleic acid, histone, amphipathic compound.  
 Disclosure; Page 34-35; 69pp; English.  
 The present sequence encodes a human neuronal thread protein AD7c-NTP. AD7c-NTP is overexpressed in brains with Alzheimer's disease at early and intermediate stages of neurodegeneration. The expression of AD7c-NTP may be reduced using the method of the invention. The specification describes a method for inducing prolonged in vivo gene expression in a mammal. The method comprises contacting a non-muscular tissue with a composition comprising a nucleic acid, histone and an amphipathic compound. The method is useful for inducing prolonged in vivo gene expression in non-muscular tissue of a mammal, e.g. neuronal tissue, central nervous system (CNS) tissue, tissue comprising a post-mitotic neuronal cell, cortical neuronal cell or hippocampal neuronal cell, glial cell, or vascular endothelial cell. The method is useful in gene therapy applications to treat Alzheimer's disease, where the composition comprises antisense AD7c-NTP nucleic acid  
 Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 1442; DB 8; Length 1442;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TTTTGTGAGATGGAGTTTTCGCTCTTGTGCGCAGGCTGAGTCAATGGCGCAA 60  
 DB 1 TTTTGTGAGATGGAGTTTTCGCTCTTGTGCGCAGGCTGAGTCAATGGCGCAA 60  
 QY 61 TCTAGCTCACCGCAACCTCCGCTCCCGGGTTCAGGCGATTCCTGCTCAGCTCC 120  
 DB 61 TCTAGCTCACCGCAACCTCCGCTCCCGGGTTCAGGCGATTCCTGCTCAGCTCC 120  
 QY 121 CAGTAGCTGGGATTACAGSCATGTGACCCAGCTCGGCTAAATTTTGTATTTT 180  
 DB 121 CAGTAGCTGGGATTACAGSCATGTGACCCAGCTCGGCTAAATTTTGTATTTT 180  
 QY 181 TAGAGATGGAGTTTCTCCATGTTGCTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC 240  
 DB 181 TAGAGATGGAGTTTCTCCATGTTGCTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC 240  
 QY 241 CCTCGCTCTCGGCTCCCAAGTGTAGATACAGACTGGCCACCAATGCGCGCTCTGCC 300  
 DB 241 CCTCGCTCTCGGCTCCCAAGTGTAGATACAGACTGGCCACCAATGCGCGCTCTGCC 300  
 QY 301 TGGCTAAATTTTGTGTTAGAAACAGGTTTCACTCATGTGCGCAAGCTGCTCTCTGAGC 360  
 DB 301 TGGCTAAATTTTGTGTTAGAAACAGGTTTCACTCATGTGCGCAAGCTGCTCTCTGAGC 360  
 QY 361 TCAAGCAGTCCACCTGCTCAGGCTCCCAAGTGTGGGATTACAGGCGTGCAGCGTGC 420  
 DB 361 TCAAGCAGTCCACCTGCTCAGGCTCCCAAGTGTGGGATTACAGGCGTGCAGCGTGC 420  
 QY 421 CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCCTCTTACCCAGGATGAAG 480  
 DB 421 CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCCTCTTACCCAGGATGAAG 480  
 QY 481 TGCAGTGGTGTGATCACAGCTCAGTGGAGCTTCACTCTCTGAGATCAAGCATCTCTCTG 540  
 DB 481 TGCAGTGGTGTGATCACAGCTCAGTGGAGCTTCACTCTCTGAGATCAAGCATCTCTCTG 540  
 QY 541 CCTCAGCTCCCAAGTAGCTGGGACAAAGACATGCACACTTACACCTGGCTAAATTTT 600

Db 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGCTAATTTT 600  
 Qy 601 TTTTATATTTTAAATTTTGGACAGAGTCTCACTCTGTCAACCGCTGAGTGCACT 660  
 Db 601 TTTTATATTTTAAATTTTGGACAGAGTCTCACTCTGTCAACCGCTGAGTGCACT 660  
 Qy 661 GCGCAATCTGGCTCAGTCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCA 720  
 Db 661 GCGCAATCTGGCTCAGTCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCA 720  
 Qy 721 GCTCTCTAGTAGCTGGGACTACAGCGCCCAACAGCTAGCTAATTTTGTATTTT 780  
 Db 721 GCTCTCTAGTAGCTGGGACTACAGCGCCCAACAGCTAGCTAATTTTGTATTTT 780  
 Qy 781 TAGTAGAGATGGGTTTCAACATGTTGCCAGGTTGATCTTGATCTCTGGACCTTGTGATC 840  
 Db 781 TAGTAGAGATGGGTTTCAACATGTTGCCAGGTTGATCTTGATCTCTGGACCTTGTGATC 840  
 Qy 841 TGCTGCTCGCTCCCAAGTGTGGGATTTACAGGCGTGAGCCACACGCGCCGCTTA 900  
 Db 841 TGCTGCTCGCTCCCAAGTGTGGGATTTACAGGCGTGAGCCACACGCGCCGCTTA 900  
 Qy 901 TTTTAAATTTTGTGTTGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGTCAAT 960  
 Db 901 TTTTAAATTTTGTGTTGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGTCAAT 960  
 Qy 961 GSCCAATCTCGCTCAGTCAACCTCTGCTCCCGGGCTCAAGGATTTCTCTGTCTCA 1020  
 Db 961 GSCCAATCTCGCTCAGTCAACCTCTGCTCCCGGGCTCAAGGATTTCTCTGTCTCA 1020  
 Qy 1021 GCTCTCAGAGCTGGGATTTAGGCACTCTGCAACACACCCGCTAATTTTGTATTT 1080  
 Db 1021 GCTCTCAGAGCTGGGATTTAGGCACTCTGCAACACACCCGCTAATTTTGTATTT 1080  
 Qy 1081 TCATTAGAGCGGGGTTTCAACATATTTGTCTAGGCTGGTCTCAAACTCTGACCTCAGT 1140  
 Db 1081 TCATTAGAGCGGGGTTTCAACATATTTGTCTAGGCTGGTCTCAAACTCTGACCTCAGT 1140  
 Qy 1141 GACCACTGCTCAGCTTCAAGTGTCTGGGATTTACAGGCTGAGCCACCTCAACCG 1200  
 Db 1141 GACCACTGCTCAGCTTCAAGTGTCTGGGATTTACAGGCTGAGCCACCTCAACCG 1200  
 Qy 1201 CCGGCTAAATTTAGATAAATAATATGTAGCAATGGGGGCTTTGCTATGTTGCCAGGCT 1260  
 Db 1201 CCGGCTAAATTTAGATAAATAATATGTAGCAATGGGGGCTTTGCTATGTTGCCAGGCT 1260  
 Qy 1261 GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACCAACCCAGCTCA 1320  
 Db 1261 GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACCAACCCAGCTCA 1320  
 Qy 1321 CATTTTAAACAGTTACATCTTTATTTTAGTATACAGAAAGTAAATACAAATGATGT 1380  
 Db 1321 CATTTTAAACAGTTACATCTTTATTTTAGTATACAGAAAGTAAATACAAATGATGT 1380  
 Qy 1381 CAAACTGCAATTCAGTAGTAAACAGAGTCTTTTATATCTTTTAAACAAAGCTTTAG 1440  
 Db 1381 CAAACTGCAATTCAGTAGTAAACAGAGTCTTTTATATCTTTTAAACAAAGCTTTAG 1440  
 Qy 1441 CA 1442  
 Db 1441 CA 1442

RESULT 5  
 AAL54224  
 ID AAL54224 standard; DNA; 1442 BP.  
 XX  
 AC AAL54224;  
 XX  
 DT 27-MAR-2003 (first entry)  
 XX  
 DE AD7C-neural thread protein encoding DNA.  
 XX

KW Nootropic; neuroprotective; cell death; tissue necrosis; NTP;  
 KW neural thread protein; neurodegenerative disorder; Alzheimer's disease;  
 XX gene; db.  
 OS Unidentified.  
 XX WO200292115-A2.  
 XX 21-NOV-2002.  
 XX 16-MAY-2002; 2002WO-CA000712.  
 XX 16-MAY-2001; 2001US-0290971P.  
 XX (NYMO-) NYMOX CORP.  
 XX Averbach PA;  
 XX WPI; 2003-129234/12.  
 XX P-PSDB; AB119445.  
 PT Preventing and/or inhibiting cell death and/or tissue necrosis in a  
 PT tissue for treating a neurodegenerative disorder, e.g. Alzheimer's  
 PT disease, by contacting the live tissue with at least one segment of  
 PT neural thread proteins (NTP).  
 XX Disclosure; Fig 1; 60pp; English.  
 XX The invention relates to a novel method for preventing and/or inhibiting  
 CC cell death and/or tissue necrosis in a tissue comprising contacting the  
 CC live tissue with at least one segment of neural thread proteins (NTP).  
 CC The methods are composition are useful for treating a neurodegenerative  
 CC disorder, such as Alzheimer's disease. This polynucleotide sequence  
 CC represents the DNA encoding the AD7C-NTP protein relating to the  
 CC invention  
 XX  
 SQ Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 1442; DB 8; Length 1442;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 TTTTATATTTTGGAGATGGAGTTTTCGCTCTTTGTTGCCAGGCTGGAGTGAATGGCGAA 60  
 Db 1 TTTTATATTTTGGAGATGGAGTTTTCGCTCTTTGTTGCCAGGCTGGAGTGAATGGCGAA 60  
 Qy 61 TCTCAGCTCAGCGCAACCTCCGCTCCCGGGTTCAAGCGATTTCTCTGCTCAGCTCCC 120  
 Db 61 TCTCAGCTCAGCGCAACCTCCGCTCCCGGGTTCAAGCGATTTCTCTGCTCAGCTCCC 120  
 Qy 121 CAGTAGCTGGGATTTACAGGCATGTGCACCCAGCTCCGCTAATTTTGTATTTTGTAG 180  
 Db 121 CAGTAGCTGGGATTTACAGGCATGTGCACCCAGCTCCGCTAATTTTGTATTTTGTAG 180  
 Qy 181 TAGAGATGGAGTTTCTCCATGTTGCTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC 240  
 Db 181 TAGAGATGGAGTTTCTCCATGTTGCTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC 240  
 Qy 241 CCTCGTCTCGGCTCCCAAGTGTAGATACAGGACTGGCCACCATGCCGGCTCTGCC 300  
 Db 241 CCTCGTCTCGGCTCCCAAGTGTAGATACAGGACTGGCCACCATGCCGGCTCTGCC 300  
 Qy 301 TGGCTAATTTTGGTGTAGAAACAGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGC 360  
 Db 301 TGGCTAATTTTGGTGTAGAAACAGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGC 360  
 Qy 361 TCAAGCAGTCCACCTGCTCCAGGCTCCCAAGTGTGGATTTACAGGCGTGCAGCGTGC 420  
 Db 361 TCAAGCAGTCCACCTGCTCCAGGCTCCCAAGTGTGGATTTACAGGCGTGCAGCGTGC 420  
 Qy 421 CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCCCACTTTACCCAGGATGAAG 480  
 Db 421 CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCCCACTTTACCCAGGATGAAG 480

481 TGCAAGTGGTGTGATCAAGCTCACTGAGCCTTCAACTCTGAGATCAAGCATCTCTCTG 540  
 481 TGCAAGTGGTGTGATCAAGCTCACTGAGCCTTCAACTCTGAGATCAAGCATCTCTCTG 540  
 541 CCTCAGCCTCCCAAGTAGCTGGACCAAGAGACATGACCACTACCTGGCTGAATTTT 600  
 541 CCTCAGCCTCCCAAGTAGCTGGACCAAGAGACATGACCACTACCTGGCTGAATTTT 600  
 601 TTTTATTTTAAATTTTGTGAGACAGAGTCTCAACTCTGTCAACAGGCTGGAGTCAAGT 660  
 601 TTTTATTTTAAATTTTGTGAGACAGAGTCTCAACTCTGTCAACAGGCTGGAGTCAAGT 660  
 661 GGGCAATCTTGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCA 720  
 661 GGGCAATCTTGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCA 720  
 721 GCTCTCTGAGTCTGGGACTACAGGGCCCAAGGCTGAGTCAATTTTGTATTTT 780  
 721 GCTCTCTGAGTCTGGGACTACAGGGCCCAAGGCTGAGTCAATTTTGTATTTT 780  
 781 TAGTAGAGTGGGGTTTCAACATCTTCCGCAAGTGTGATCTTGTACCTGAGTGTGATC 840  
 781 TAGTAGAGTGGGGTTTCAACATCTTCCGCAAGTGTGATCTTGTACCTGAGTGTGATC 840  
 841 TGCCTGCTCGGCTCCCAAGTGTGGGATTAACAGGCTGAGCCACACGCGCGCTT 900  
 841 TGCCTGCTCGGCTCCCAAGTGTGGGATTAACAGGCTGAGCCACACGCGCGCTT 900  
 901 TTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTACCCAGGCTGGAGTCAAT 960  
 901 TTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTACCCAGGCTGGAGTCAAT 960  
 961 GGGCAATCTGCTCACTGCAACCTCTGCTCCCGGGTTCAAGGATTTCTCTGCTCA 1020  
 961 GGGCAATCTGCTCACTGCAACCTCTGCTCCCGGGTTCAAGGATTTCTCTGCTCA 1020  
 1021 GCTTCCCAAGCAGTGGGATTAAGGCACTGCGCAACACACCCCGCTAAATTTTGTATTT 1080  
 1021 GCTTCCCAAGCAGTGGGATTAAGGCACTGCGCAACACACCCCGCTAAATTTTGTATTT 1080  
 1081 TCATTAGAGCGGGGTTTCAACATTTTGTAGGCTGGTCTCAACTCTGACCTCAGT 1140  
 1081 TCATTAGAGCGGGGTTTCAACATTTTGTAGGCTGGTCTCAACTCTGACCTCAGT 1140  
 1141 GACCCACCTGCTCAGCTTCCAAAGTGTGGGATTAACAGGCTGAGCCACCTCACCAG 1200  
 1141 GACCCACCTGCTCAGCTTCCAAAGTGTGGGATTAACAGGCTGAGCCACCTCACCAG 1200  
 1201 CCGGCTAAATTTAGATAAAAAATATAGCAATGGGGGCTTTGCTATGTTGCCAGGCT 1260  
 1201 CCGGCTAAATTTAGATAAAAAATATAGCAATGGGGGCTTTGCTATGTTGCCAGGCT 1260  
 1261 GGTCTCAAACTTCTGCTTCAATGCAATCTTCAATAGGCAACACACCCAGCAATCA 1320  
 1261 GGTCTCAAACTTCTGCTTCAATGCAATCTTCAATAGGCAACACACCCAGCAATCA 1320  
 1321 CATTTTAAACAGTTTACATCTTTATTTAGTATCTAGAAAGTAAATCAATAAATCT 1380  
 1321 CATTTTAAACAGTTTACATCTTTATTTAGTATCTAGAAAGTAAATCAATAAATCT 1380  
 1381 CAACTGCAATTTCACTAGTAAACAGAGTCTTTTATTAATCTTTTAAACAAAGCTTTAG 1440  
 1381 CAACTGCAATTTCACTAGTAAACAGAGTCTTTTATTAATCTTTTAAACAAAGCTTTAG 1440  
 1441 CA 1442  
 1441 CA 1442

RESULT 6  
 ADB37519  
 ID ADB37519 standard; DNA; 1442 BP.

XX AC ADB37519;  
 XX DT 04-DEC-2003 (first entry)  
 XX DE Human neural thread protein AD7c-NTP coding sequence.  
 XX KW Cytostatic; Antitumour; Antipsoriatic; Dermatological;  
 KW Antiatherosclerotic; Antiarteriosclerotic; Vasotropic;  
 KW Immunosuppressive; Tranquillizer; Antileptic; Antineoplastic;  
 KW neural thread protein; neuritic sprouting; gene; ds.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT CDS 15..1142  
 FT /tag= a  
 FT /product= "AD7c-NTP"  
 XX PN WO2003008444-A2.  
 XX PD 30-JAN-2003.  
 XX PF 19-JUL-2002; 2002WO-CA001106.  
 XX PR 19-JUL-2001; 2001US-0306150P.  
 PR 19-JUL-2001; 2001US-0306161P.  
 PR 16-NOV-2001; 2001US-0331477P.  
 XX PA (NIMO-) NYMOX CORP.  
 XX PI Averbach PA, Gemmell J;  
 XX DR WPI: 2003-248000/24.  
 DR P-PSDB; ADB37642.  
 FT Novel Related peptide or AD7c-neural thread peptide, useful for treating  
 PT unwanted cellular proliferations, glandular hyperplasia, unwanted facial  
 PT hair, warts and unwanted fatty tissue.  
 XX Disclosure; Fig 1; 109pp; English.  
 XX The present invention relates to AD7c-neural thread protein (NTP) and  
 CC related proteins and peptides (I; ADB37528-ADB37641). The sequences are  
 CC useful for treating a condition in a patient requiring removal or  
 CC destruction of cells. The condition can be selected from benign or  
 CC malignant tumour of a tissue, hyperplasia, hypertrophy or overgrowth of a  
 CC tissue, virally, bacterially or parasitically altered tissue, or  
 CC malformation of a tissue, where the tissue is selected from lung, breast,  
 CC stomach, pancreas, prostate, bladder, bone, ovary, skin, kidney, salivary  
 CC colon, intestine, stomach, rectum, oesophagus, heart, spleen, salivary  
 CC gland, blood, brain and its coverings, spinal cord and its coverings,  
 CC muscle, connective tissue, adrenal, parathyroid, thyroid, uterus, testis,  
 CC pituitary, reproductive organs, liver, gall bladder, eye, ear, nose,  
 CC throat, tonsils, mouth, lymph nodes and lymphoid tissue. The condition is  
 CC preferably tonsillary hypertrophy, prostatic hyperplasia, psoriasis,  
 CC eczema, dermatosis, cosmetic modification to a tissue (skin, eye, ear,  
 CC nose, throat, mouth, muscle, connective, hair or breast tissue), vascular  
 CC disease (atherosclerosis or arteriosclerosis), haemorrhoids, varicose  
 CC veins, inflammatory disease, autoimmune disease, metabolic disease,  
 CC hereditary/genetic disease, traumatic disease or physical injury,  
 CC nutritional deficiency disease, infectious disease, amyloid disease,  
 CC fibrosis disease, storage disease, congenital malformation, enzyme  
 CC deficiency disease, poisoning, intoxication, environmental disease,  
 CC radiation disease, endocrine disease, degenerative disease and mechanical  
 CC disease. The peptides are useful for treating unwanted cellular  
 CC proliferations, glandular (e.g. prostate) hyperplasia, unwanted facial  
 CC hair, warts and unwanted fatty tissue, or for preparing antibodies that  
 CC recognize and/or bind to Related proteins. Related peptides or NTP  
 CC peptides. The present sequence is the coding sequence for AD7c-NTP, used  
 CC to illustrate the invention. AD7c-NTP is a ~41kd membrane associated  
 CC phosphoprotein with functions associated with neuritic sprouting.  
 XX

SQ Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 U; 0 Other;									
Query Match 100.0%; Score 1442; DB 10; Length 1442;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	TTTTTTTTTTCAGATGGAGTTTTCGCTCTGTTGCCACAGGCTGGAGTGCATGGCGAA	60						
DB	1	TTTTTTTTTTCAGATGGAGTTTTCGCTCTGTTGCCACAGGCTGGAGTGCATGGCGAA	60						
QY	61	TCTCAGCTCACGCAACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCTCAGCTCC	120						
DB	61	TCTCAGCTCACGCAACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCTCAGCTCC	120						
QY	121	CAGTAGCTGGATTCAGAGCATGTGACCCAGCTCGGCTAAATTTTGTATTTTTTTAG	180						
DB	121	CAGTAGCTGGATTCAGAGCATGTGACCCAGCTCGGCTAAATTTTGTATTTTTTTAG	180						
QY	181	TAGAGATGAGTTTCTCCATGTTTGTTCAGGCTGGTCTCGAACTCCCGACCTCAGATGTC	240						
DB	181	TAGAGATGAGTTTCTCCATGTTTGTTCAGGCTGGTCTCGAACTCCCGACCTCAGATGTC	240						
QY	241	CCTCGCTCTCGGCTCCCAAGTGTCTAGATACAGGACTGGCCACCATGCCCCGCTCTGCC	300						
DB	241	CCTCGCTCTCGGCTCCCAAGTGTCTAGATACAGGACTGGCCACCATGCCCCGCTCTGCC	300						
QY	301	TGCTAAATTTTGTGTAGAAACAGGTTTCTCTGATGTGCCCAAGCTGGTCTCTGAGC	360						
DB	301	TGCTAAATTTTGTGTAGAAACAGGTTTCTCTGATGTGCCCAAGCTGGTCTCTGAGC	360						
QY	361	TCAAGCAGTCCACCTCGCTCAGCTCCCAAGTGTCTGGATTACAGGCGTGACCGTGC	420						
DB	361	TCAAGCAGTCCACCTCGCTCAGCTCCCAAGTGTCTGGATTACAGGCGTGACCGTGC	420						
QY	421	CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG	480						
DB	421	CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG	480						
QY	481	TGCAGTGTGTATTCACAGCTCAGCTCAGGCTTCAACTCTCAGATCAAGATCTCTCTG	540						
DB	481	TGCAGTGTGTATTCACAGCTCAGCTCAGGCTTCAACTCTCAGATCAAGATCTCTCTG	540						
QY	541	CCTCAGCTCCCAAGTAGCTGGACCAAGACATGCACCACTACACCTGGCTAAATTTTFA	600						
DB	541	CCTCAGCTCCCAAGTAGCTGGACCAAGACATGCACCACTACACCTGGCTAAATTTTFA	600						
QY	601	TTTTTATTTTAAATTTTGTGACACAGGTCTCAACTCTGTCAACCCAGGCTGGAGTGCAGT	660						
DB	601	TTTTTATTTTAAATTTTGTGACACAGGTCTCAACTCTGTCAACCCAGGCTGGAGTGCAGT	660						
QY	661	GGCGCAATCTGGCTCAGTGCACCTCTGCTCCCGGGTTCAAGTTATTTCTCGGCCCA	720						
DB	661	GGCGCAATCTGGCTCAGTGCACCTCTGCTCCCGGGTTCAAGTTATTTCTCGGCCCA	720						
QY	721	GCTCTGTAGTAGCTGGAGTACAGCGCCCAACAGCTAGTCAATTTTTTTGTATTTT	780						
DB	721	GCTCTGTAGTAGCTGGAGTACAGCGCCCAACAGCTAGTCAATTTTTTTGTATTTT	780						
QY	781	TAGTAGAGATGGGGTTTACCATGTTCCGACAGGTTGATCTTGATCTCGGACCTTGTGATC	840						
DB	781	TAGTAGAGATGGGGTTTACCATGTTCCGACAGGTTGATCTTGATCTCGGACCTTGTGATC	840						
QY	841	TGCTCGCTCGGCTCCCAAGTGTCTGGATTACAGGCGTGAGCCACAGCGCCGCTTFA	900						
DB	841	TGCTCGCTCGGCTCCCAAGTGTCTGGATTACAGGCGTGAGCCACAGCGCCGCTTFA	900						
QY	901	TTTTTAAATTTTGTGTTTGAATGGAATCTCAGTCTCTGTACCCAGGCTGGAGTGCAT	960						
DB	901	TTTTTAAATTTTGTGTTTGAATGGAATCTCAGTCTCTGTACCCAGGCTGGAGTGCAT	960						
QY	961	GGCCAAATCTCGGCTCAGTGCACCTCTGCTCCCGGGCTCAAGCGATTCTCTGCTCA	1020						
DB	961	GGCCAAATCTCGGCTCAGTGCACCTCTGCTCCCGGGCTCAAGCGATTCTCTGCTCA	1020						

QY	1021	GCCTCCCAAGCAGCTGGGATTAACGGCACACCTGGCCACACACCCCGCTAATTTTGTATTT	1080
DB	1021	GCCTCCCAAGCAGCTGGGATTAACGGCACACCTGGCCACACACCCCGCTAATTTTGTATTT	1080
QY	1081	TCATTTAGAGGGGGGTTTTCACCATATTTTGTTCAGGCTGGTCTCAAACTCCTGACCTCAGGT	1140
DB	1081	TCATTTAGAGGGGGGTTTTCACCATATTTTGTTCAGGCTGGTCTCAAACTCCTGACCTCAGGT	1140
QY	1141	GACCCACCTGCTCAGCTTCCAAAGTGTCTGGGATTTACAGGCGTGAGCCACTCACCCAG	1200
DB	1141	GACCCACCTGCTCAGCTTCCAAAGTGTCTGGGATTTACAGGCGTGAGCCACTCACCCAG	1200
QY	1201	CGGGCTAATTTAGATAAAAAAATATGTAGCAATGSGGGGCTTCTGTATGTTGCCAGGCT	1260
DB	1201	CGGGCTAATTTAGATAAAAAAATATGTAGCAATGSGGGGCTTCTGTATGTTGCCAGGCT	1260
QY	1261	GGTCTCAAACTTCTGGCTTCATGCAATCCTTCCAAATGAGCCACAAACCCAGCCAGTCA	1320
DB	1261	GGTCTCAAACTTCTGGCTTCATGCAATCCTTCCAAATGAGCCACAAACCCAGCCAGTCA	1320
QY	1321	CATTTTAAACAGTTACATCTTTTATTTTAGTATATCTAGAAAAGTAAATCAATAAATGT	1380
DB	1321	CATTTTAAACAGTTACATCTTTTATTTTAGTATATCTAGAAAAGTAAATCAATAAATGT	1380
QY	1381	CAAACTGCAAAATTCAGTAGTAACAGAGTCTTTTTTATAACTTTTTAAACAAAGCTTTAGAG	1440
DB	1381	CAAACTGCAAAATTCAGTAGTAACAGAGTCTTTTTTATAACTTTTTAAACAAAGCTTTAGAG	1440
QY	1441	CA 1442	
DB	1441	CA 1442	
RESULT 7			
ACC84453			
ID	ACC84453 standard; DNA; 1442 BP.		
AC	ACC84453;		
XX			
DT	28-AUG-2003 (first entry)		
XX	AD7c-NTP encoding sequence.		
DE	DE		
XX	Cytostatic; Antibacterial; Immunosuppressive; Antiinflammatory;		
KW	neural thread protein; NTP; tumour; ds.		
XX	Unidentified.		
XX	WO2003008443-A2.		
PN	30-JAN-2003.		
PD	19-JUL-2002; 2002WO-CA001105.		
PF	19-JUL-2001; 2001US-0306150P.		
PR	19-JUL-2001; 2001US-0306161P.		
PR	16-NOV-2001; 2001US-0331477P.		
XX	(NYMO-) NYMOX CORP.		
PA	Averback PA;		
XX	WPI; 2003-247999/24.		
DR	Novel neural thread protein peptide, referred as cell death peptide,		
PT	useful for treating prostatic hyperplasia, psoriasis, eczema, dermatosis,		
PT	atherosclerosis, cosmetic modification to skin, throat, mouth, muscle.		
XX	Disclosure; Fig 1; 77pp; English.		
PS	The present invention relates to a neural thread protein (NTP) peptide		
CC	referred to as cell death peptide. Thought to be cytostatic,		



CC antibacterial, immunosuppressive and antiinflammatory. It is useful for  
 CC treating a condition in a patient requiring removal or destruction of  
 CC cells, for treating a condition such as benign or malignant tumor,  
 CC inflammatory disease, autoimmune disease and infectious disease. The  
 CC peptide useful for treatment is derived from the amino acid sequence for  
 CC a pancreatic thread protein. The peptide is conjugated, linked or bound  
 CC to a molecule chosen from antibody or its fragment, antibody-like binding  
 CC molecule, where the molecule has a higher affinity for binding to a tumor  
 CC or other target than binding to other cells. Treatment using NTP peptides  
 CC can remove benign tumors with less risk and fewer of the undesirable side  
 CC effects of surgery. The present sequence is an NTP encoding sequence  
 XX  
 XX  
 SQ Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 U; 0 Other;

Query Match		100.0%; Score 1442; DB 10; Length 1442;
Best Local Similarity		100.0%; Pred. No. 0;
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	TTTTTTTTTTGAGATGGAGTTTTCGCTCTGTGTGTCGCCAGGCTGGAGTGCATATGGCGCAA 60
DB	1	TTTTTTTTTTGAGATGGAGTTTTCGCTCTGTGTGTCGCCAGGCTGGAGTGCATATGGCGCAA 60
QY	61	TCTCAGCTCAGCGCAACCTCCGCTCCCGGTTTCAAGCGATTTCTCTGCTCAGCTCAGCTCCC 120
DB	61	TCTCAGCTCAGCGCAACCTCCGCTCCCGGTTTCAAGCGATTTCTCTGCTCAGCTCAGCTCCC 120
QY	121	CAGTAGCTGGGATTACAGGATGTGACCCAGCTCCGCTTAATTTGTATTTTTTTTTTAG 180
DB	121	CAGTAGCTGGGATTACAGGATGTGACCCAGCTCCGCTTAATTTGTATTTTTTTTTTAG 180
QY	181	TAGAGATGAGTTTCTCCATTTGGTTCAGGCTGGTCTCGAATCTCCGACCTCAGATGATC 240
DB	181	TAGAGATGAGTTTCTCCATTTGGTTCAGGCTGGTCTCGAATCTCCGACCTCAGATGATC 240
QY	241	CTCCGCTCTCGGCTCCGAAAGTGTAGATACAGGATGGCCACCAATGCGCGCTCTGCC 300
DB	241	CTCCGCTCTCGGCTCCGAAAGTGTAGATACAGGATGGCCACCAATGCGCGCTCTGCC 300
QY	301	TGCTTAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGCTCTCTGAGC 360
DB	301	TGCTTAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGCTCTCTGAGC 360
QY	361	TCAGCAGTCCACCTCGCTCAGCTCCAAAGTGTGGGATTCAGGCGTGCAGCGCTGC 420
DB	361	TCAGCAGTCCACCTCGCTCAGCTCCAAAGTGTGGGATTCAGGCGTGCAGCGCTGC 420
QY	421	CTGGCCCTTTTATTTATTTTATTTTAAAGACAGAGTGTCCCACTTTACCCAGGATGA 480
DB	421	CTGGCCCTTTTATTTATTTTATTTTAAAGACAGAGTGTCCCACTTTTACCCAGGATGA 480
QY	481	TGCAAGTGTGTATCACAGCTCACTGAGCTTCACTGAGCTTCACTGAGATCAAGCATCTCTG 540
DB	481	TGCAAGTGTGTATCACAGCTCACTGAGCTTCACTGAGCTTCACTGAGATCAAGCATCTCTG 540
QY	541	CCTCAGCTCCCAAGTAGTGGGACCAAGACATGCACCACTACACCTCGCTTAATTTT 600
DB	541	CCTCAGCTCCCAAGTAGTGGGACCAAGACATGCACCACTACACCTCGCTTAATTTT 600
QY	601	TTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660
DB	601	TTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660
QY	661	GGCGCAATCTGGCTCAGTCACTGAGCTTCCGCTCCCGGTTTCAAGTATTTCTCTGCCCA 720
DB	661	GGCGCAATCTGGCTCAGTCACTGAGCTTCCGCTCCCGGTTTCAAGTATTTCTCTGCCCA 720
QY	721	GCTCTCTGAGTGTGGGACTACAGCGGCCACCAAGCTAGCTTAATTTTATTTTATTTT 780
DB	721	GCTCTCTGAGTGTGGGACTACAGCGGCCACCAAGCTAGCTTAATTTTATTTTATTTT 780
QY	781	TAGTAGAGATGGGTTTCACTGAGTGTGGGATTTGATCTCTGACCTCTGATC 840
DB	781	TAGTAGAGATGGGTTTCACTGAGTGTGGGATTTGATCTCTGACCTCTGATC 840

QY	841	TGCTGCTCGCTCGGCTCCCAAGTGTGGGATTCAGGCGTGCAGCCACCGCCGCTTA 900	
DB	841	TGCTGCTCGCTCGGCTCCCAAGTGTGGGATTCAGGCGTGCAGCCACCGCCGCTTA 900	
QY	901	TTTTTAATTTTGTGTGTGTTGTTGAAATCTCACTCTCTTATCCAGGCTGGAGTGCAT 960	
DB	901	TTTTTAATTTTGTGTGTGTTGTTGAAATCTCACTCTCTTATCCAGGCTGGAGTGCAT 960	
QY	961	GGCCAAATCTCGGCTCAGTGCACCTCTGCTCCCGGCTCAAGCGATTTCTCTGCTCA 1020	
DB	961	GGCCAAATCTCGGCTCAGTGCACCTCTGCTCCCGGCTCAAGCGATTTCTCTGCTCA 1020	
QY	1021	GCTCTCCCAAGCAGCTGGGATTCAGGCGCTGCACCAACCCCGCTTAATTTTGTATTT 1080	
DB	1021	GCTCTCCCAAGCAGCTGGGATTCAGGCGCTGCACCAACCCCGCTTAATTTTGTATTT 1080	
QY	1081	TCAATTAGAGCGGGGTTTCCATATTTGTTCAGGCTGGTCTCAAACTCTCTGACCTCAGGT 1140	
DB	1081	TCAATTAGAGCGGGGTTTCCATATTTGTTCAGGCTGGTCTCAAACTCTCTGACCTCAGGT 1140	
QY	1141	GACCCACCTGCTCAGCTTCCCAAGTGTGGGATTCAGGCGTGCAGCCACCTCAGCCAG 1200	
DB	1141	GACCCACCTGCTCAGCTTCCCAAGTGTGGGATTCAGGCGTGCAGCCACCTCAGCCAG 1200	
QY	1201	CGGCTTAATTTAGATAAAAAAATATGTAGCAATGGGGGTCTTGTATGTTGCCAGGCT 1260	
DB	1201	CGGCTTAATTTAGATAAAAAAATATGTAGCAATGGGGGTCTTGTATGTTGCCAGGCT 1260	
QY	1261	GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACCAACCCAGCCAGTCA 1320	
DB	1261	GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACCAACCCAGCCAGTCA 1320	
QY	1321	CATTTTTTAAACAGTTACATCTTTATTTAGTATATCTAGAAAGTAAATACATAAATGTT 1380	
DB	1321	CATTTTTTAAACAGTTACATCTTTATTTAGTATATCTAGAAAGTAAATACATAAATGTT 1380	
QY	1381	CAAACTGCAATTCAGTAGTAAACAGAGTCTTTTATTAACCTTTTAAACAAAGCTTTAGAG 1440	
DB	1381	CAAACTGCAATTCAGTAGTAAACAGAGTCTTTTATTAACCTTTTAAACAAAGCTTTAGAG 1440	
QY	1441	CA 1442	
DB	1441	CA 1442	
RESULT 8			
ADRI14408			
ID	ADRI14408 standard; DNA; 1442 BP.		
XX	ADRI14408;		
XX	21-OCT-2004 (first entry)		
XX	Human NF-kappaB pathway-associated gene SeqID409.		
DE	NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;		
KW	antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;		
KW	antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;		
KW	immunosuppressive; vulnery; gene therapy; immune disorder;		
KW	inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;		
KW	hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;		
KW	hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;		
KW	X-linked anhidrotic ectodermal dysplasia; immunodeficiency;		
KW	viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;		
KW	rheumatoid arthritis; host cell survival; evasion of immune response;		
KW	autoimmune disorder; hyper immune activity;		
KW	aberrant acute phase response; hypercongenital condition; birth defect;		
KW	necrotic lesion; wound; organ transplant rejection;		
KW	aberrant signal transduction; proliferating disorder; cancer;		
KW	HIV propagation; gene; ds; human.		



XX Homo sapiens.  
 XX WO2004065577-A2.  
 XX PD 05-AUG-2004.  
 XX 13-JAN-2004; 2004WO-US000798.  
 XX 14-JAN-2003; 2003US-0440068P.  
 XX PR 12-MAY-2003; 2003US-0469575P.  
 XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX Nadler SG, Neubauer MG, Feder JN, Carman J;  
 XX WPI; 2004-562168/54.  
 XX DR P-PSDB; ADR14409.  
 XX New isolated polynucleotides and polypeptides associated with NF-kappaB  
 PT pathway, useful for diagnosing, treating, or preventing disorders or  
 PT diseases associated with NF-kappaB pathway.  
 XX Claim 1; SEQ ID NO 409; 237pp; English.  
 CC This invention relates to the novel association of protein sequences (and  
 CC the genes which encode them) to the NF-kappaB pathway. The invention may  
 CC be useful for the production of compounds with an anti-inflammatory,  
 CC cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,  
 CC gastrointestinal-Gen, antiaesthetic, antiarteriosclerotic,  
 CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or  
 CC vulnery activity or for gene therapy. The proteins and nucleotides are  
 CC useful for diagnosing, preventing, treating, or ameliorating conditions  
 CC or diseases associated with the NF-kappaB pathway. The condition is an  
 CC immune disorder, an inflammatory disorder, an inflammatory disorder  
 CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,  
 CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM  
 CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic  
 CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,  
 CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell  
 CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory  
 CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick  
 CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper  
 CC immune activity, disorders related to aberrant acute phase responses,  
 CC hypergenital conditions, birth defects, necrotic lesions, wounds,  
 CC organ transplant rejection, conditions related to organ transplant  
 CC rejection, disorders related to aberrant signal transduction,  
 CC proliferating disorders, cancers and HIV propagation in cells infected  
 CC with other viruses. The present sequence is that of a human gene which is  
 CC subject to the novel association with the NF-kappaB pathway of the  
 CC invention. Note: This sequence does not appear in the specification but  
 CC was obtained by the indexer from Genbank.  
 XX SQ Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 U; 0 Other;

181	TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC	240
241	CCTCCGTTCTGGGCTCCAAAAGTCTAGATACAGGACTGGGCAACATGCCCGGCTCTGCC	300
241	CCTCCGTTCTGGGCTCCAAAAGTCTAGATACAGGACTGGGCAACATGCCCGGCTCTGCC	300
301	TGGCTAAATTTTGGTGTAGAAACAGGGTTTCACTGATGTGCCACAGCTGGTCTCCTGAGC	360
301	TGGCTAAATTTTGGTGTAGAAACAGGGTTTCACTGATGTGCCACAGCTGGTCTCCTGAGC	360
361	TCAAGCAGTCCACCTGCTCCAAAAGTGTGGGATTTACAGGCGTGCAGCGTGC	420
361	TCAAGCAGTCCACCTGCTCCAAAAGTGTGGGATTTACAGGCGTGCAGCGTGC	420
421	CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	480
421	CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	480
481	TGCAGTGTGTGATCAGAGCTCACTGAGCTTCAACTCTCTGAGATCAAGCATCTCTCTG	540
481	TGCAGTGTGTGATCAGAGCTCACTGAGCTTCAACTCTCTGAGATCAAGCATCTCTCTG	540
541	CCTCAGCTCCCAAGTAGCTGGGACCAAGACATGCACACTACACCTGGCTAAATTTTAA	600
541	CCTCAGCTCCCAAGTAGCTGGGACCAAGACATGCACACTACACCTGGCTAAATTTTAA	600
601	TTTTTATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	660
601	TTTTTATTTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	660
661	GGCGCAATTTGGCTCACTGCAACCTCTGCTCCGGGTTCAAGTTATTTCTCTGCCCCA	720
661	GGCGCAATTTGGCTCACTGCAACCTCTGCTCCGGGTTCAAGTTATTTCTCTGCCCCA	720
721	GCCTCTCAGTAGCTGGGACCTACAGGCGCCCAACACGCTAGCTAAATTTTATTTT	780
721	GCCTCTCAGTAGCTGGGACCTACAGGCGCCCAACACGCTAGCTAAATTTTATTTT	780
781	TAGTAGAGATGGGTTTCAACATGTTTGGCCAGTTGATCTTGATCTCTGGACCTTGATC	840
781	TAGTAGAGATGGGTTTCAACATGTTTGGCCAGTTGATCTTGATCTCTGGACCTTGATC	840
841	TGCGCTCTCGGCTCCCAAGAGTGTGGGATTTACAGGCGTGAGCCACCGCCGGCTTA	900
841	TGCGCTCTCGGCTCCCAAGAGTGTGGGATTTACAGGCGTGAGCCACCGCCGGCTTA	900
901	TTTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTTACCCAGGCTGGAGTCAAT	960
901	TTTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTTACCCAGGCTGGAGTCAAT	960
961	GGCCAAATCTCGGCTCACTGCAACCTCTGCTCCGGGCTCAAGGATTTCTCTCTCA	1020
961	GGCCAAATCTCGGCTCACTGCAACCTCTGCTCCGGGCTCAAGGATTTCTCTCTCA	1020
1021	GCCTCCCAAGCAGCTGGGATTTACGGGCACTCCCAACACCGCTAAATTTTGTATTT	1080
1021	GCCTCCCAAGCAGCTGGGATTTACGGGCACTCCCAACACCGCTAAATTTTGTATTT	1080
1081	TCATTTAGAGCGGGGTTTCAACATTTTGTGAGGCTGGTCTCAAACTCTGACCTCAGT	1140
1081	TCATTTAGAGCGGGGTTTCAACATTTTGTGAGGCTGGTCTCAAACTCTGACCTCAGT	1140
1141	GACCCACTGCTCAGCTTCCAAAAGTGTGGGATTTACAGGCGTGAGCCACCTCACCAG	1200
1141	GACCCACTGCTCAGCTTCCAAAAGTGTGGGATTTACAGGCGTGAGCCACCTCACCAG	1200
1201	CGGCTAAATTTAGATTAATAAATATGATGAGCAATGGGGGCTTTGCTATGTTGCCAGGCT	1260
1201	CGGCTAAATTTAGATTAATAAATATGATGAGCAATGGGGGCTTTGCTATGTTGCCAGGCT	1260
1261	GGTCTCAAACTTTCTGGCTTTCAATGCAATCTTCCAAATGAGGCCAACACCCAGCAGTCA	1320
1261	GGTCTCAAACTTTCTGGCTTTCAATGCAATCTTCCAAATGAGGCCAACACCCAGCAGTCA	1320

Query Match	100.0%;	Score 1442;	DB 13;	Length 1442;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1442;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	TTTTTTTTTTTTCAGATGGAGTTTTCGCTCTGCTGTCGCCAGGCTGGAGTGCATGGCGCAA	60		
DB 1	TTTTTTTTTTTTCAGATGGAGTTTTCGCTCTGCTGTCGCCAGGCTGGAGTGCATGGCGCAA	60		
QY 61	TCTCAGCTCACCGCAACTCCGCTCCCGGGTTCAAGCGATTTCTCTGCTCAGCTCCC	120		
DB 61	TCTCAGCTCACCGCAACTCCGCTCCCGGGTTCAAGCGATTTCTCTGCTCAGCTCCC	120		
QY 121	CAGTACCTGGGATTAAGCATGTGCAACCGCTCGGCTAAATTTTGTATTTTATTTAG	180		
DB 121	CAGTACCTGGGATTAAGCATGTGCAACCGCTCGGCTAAATTTTGTATTTTATTTAG	180		
QY 181	TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC	240		

Qy	1321	CA T T T T T T A A C A G T T A C A T C T T T T T T A G T A T A C T A G A A G T A T A C A A T A A C A G T	1380
Db	1321	CA T T T T T T T A A C A G T T A C A T C T T T T T T T A G T A T A C T A G A A G T A T A C A A T A A C A G T	1380
Qy	1381	CA A A C C T G C A A A T T C A G T A G T A A C A G A G T C T T T T T A T A A C T T T T T T A A C A A A G C T T T T A G	1440
Db	1381	CA A A C C T G C A A A T T C A G T A G T A A C A G A G T C T T T T T A T A A C T T T T T T A A C A A A G C T T T T A G	1440
Qy	1441	CA	1442
Db	1441	CA	1442

## RESULT 9

AAT27738

AAT27738  
ID AAT27738 standard: DNA: 1418 BP.

AC AAT

13-NOV-1996 (first entry)

xx Neural thread protein coding sequence.

Neural thread protein; NTP; diagnosis; detection; Alzheimer's disease;  
 neuroectodermal tumour; malignant astrocytoma; monoclonal antibody;  
 binding fragment; ds.

OS Homo sapiens.

PH	Key	Location/Qualifiers
PH 1	Key 1	Location 1/Qualifiers 1
PH 2	Key 2	Location 2/Qualifiers 2
PH 3	Key 3	Location 3/Qualifiers 3
PH 4	Key 4	Location 4/Qualifiers 4
PH 5	Key 5	Location 5/Qualifiers 5
PH 6	Key 6	Location 6/Qualifiers 6
PH 7	Key 7	Location 7/Qualifiers 7
PH 8	Key 8	Location 8/Qualifiers 8
PH 9	Key 9	Location 9/Qualifiers 9
PH 10	Key 10	Location 10/Qualifiers 10
PH 11	Key 11	Location 11/Qualifiers 11
PH 12	Key 12	Location 12/Qualifiers 12
PH 13	Key 13	Location 13/Qualifiers 13
PH 14	Key 14	Location 14/Qualifiers 14
PH 15	Key 15	Location 15/Qualifiers 15
PH 16	Key 16	Location 16/Qualifiers 16
PH 17	Key 17	Location 17/Qualifiers 17
PH 18	Key 18	Location 18/Qualifiers 18
PH 19	Key 19	Location 19/Qualifiers 19
PH 20	Key 20	Location 20/Qualifiers 20
PH 21	Key 21	Location 21/Qualifiers 21
PH 22	Key 22	Location 22/Qualifiers 22
PH 23	Key 23	Location 23/Qualifiers 23
PH 24	Key 24	Location 24/Qualifiers 24
PH 25	Key 25	Location 25/Qualifiers 25
PH 26	Key 26	Location 26/Qualifiers 26
PH 27	Key 27	Location 27/Qualifiers 27
PH 28	Key 28	Location 28/Qualifiers 28
PH 29	Key 29	Location 29/Qualifiers 29
PH 30	Key 30	Location 30/Qualifiers 30
PH 31	Key 31	Location 31/Qualifiers 31
PH 32	Key 32	Location 32/Qualifiers 32
PH 33	Key 33	Location 33/Qualifiers 33
PH 34	Key 34	Location 34/Qualifiers 34
PH 35	Key 35	Location 35/Qualifiers 35
PH 36	Key 36	Location 36/Qualifiers 36
PH 37	Key 37	Location 37/Qualifiers 37
PH 38	Key 38	Location 38/Qualifiers 38
PH 39	Key 39	Location 39/Qualifiers 39
PH 40	Key 40	Location 40/Qualifiers 40
PH 41	Key 41	Location 41/Qualifiers 41
PH 42	Key 42	Location 42/Qualifiers 42
PH 43	Key 43	Location 43/Qualifiers 43
PH 44	Key 44	Location 44/Qualifiers 44
PH 45	Key 45	Location 45/Qualifiers 45
PH 46	Key 46	Location 46/Qualifiers 46
PH 47	Key 47	Location 47/Qualifiers 47
PH 48	Key 48	Location 48/Qualifiers 48
PH 49	Key 49	Location 49/Qualifiers 49
PH 50	Key 50	Location 50/Qualifiers 50
PH 51	Key 51	Location 51/Qualifiers 51
PH 52	Key 52	Location 52/Qualifiers 52
PH 53	Key 53	Location 53/Qualifiers 53
PH 54	Key 54	Location 54/Qualifiers 54
PH 55	Key 55	Location 55/Qualifiers 55
PH 56	Key 56	Location 56/Qualifiers 56
PH 57	Key 57	Location 57/Qualifiers 57
PH 58	Key 58	Location 58/Qualifiers 58
PH 59	Key 59	Location 59/Qualifiers 59
PH 60	Key 60	Location 60/Qualifiers 60
PH 61	Key 61	Location 61/Qualifiers 61
PH 62	Key 62	Location 62/Qualifiers 62
PH 63	Key 63	Location 63/Qualifiers 63
PH 64	Key 64	Location 64/Qualifiers 64
PH 65	Key 65	Location 65/Qualifiers 65
PH 66	Key 66	Location 66/Qualifiers 66
PH 67	Key 67	Location 67/Qualifiers 67
PH 68	Key 68	Location 68/Qualifiers 68
PH 69	Key 69	Location 69/Qualifiers 69
PH 70	Key 70	Location 70/Qualifiers 70
PH 71	Key 71	Location 71/Qualifiers 71
PH 72	Key 72	Location 72/Qualifiers 72
PH 73	Key 73	Location 73/Qualifiers 73
PH 74	Key 74	Location 74/Qualifiers 74
PH 75	Key 75	Location 75/Qualifiers 75
PH 76	Key 76	Location 76/Qualifiers 76
PH 77	Key 77	Location 77/Qualifiers 77
PH 78	Key 78	Location 78/Qualifiers 78
PH 79	Key 79	Location 79/Qualifiers 79
PH 80	Key 80	Location 80/Qualifiers 80
PH 81	Key 81	Location 81/Qualifiers 81
PH 82	Key 82	Location 82/Qualifiers 82
PH 83	Key 83	Location 83/Qualifiers 83
PH 84	Key 84	Location 84/Qualifiers 84
PH 85	Key 85	Location 85/Qualifiers 85
PH 86	Key 86	Location 86/Qualifiers 86
PH 87	Key 87	Location 87/Qualifiers 87
PH 88	Key 88	Location 88/Qualifiers 88
PH 89	Key 89	Location 89/Qualifiers 89
PH 90	Key 90	Location 90/Qualifiers 90
PH 91	Key 91	Location 91/Qualifiers 91
PH 92	Key 92	Location 92/Qualifiers 92
PH 93	Key 93	Location 93/Qualifiers 93
PH 94	Key 94	Location 94/Qualifiers 94
PH 95	Key 95	Location 95/Qualifiers 95

FT	CDS	Location/
14.	.1207	

```
FT      / *tag= a
FT      /product= "Neural thread protein."
```

PN WO9615272-A1.

23-MAY-1996.

14-NOV-1995; 95WO-US017111.

PR 14-NOV-1994: 94US-00340426.

PA (GEHO ) GEN HOSPITAL CORP.

PI De La Monte S, Wands JR;

DR WPI; 1996-259865/26.

DR P-PSDB; AAR95913.

PT Detection of neural thread protein in diagnosis of Alzheimer's disease -  
PT also NTP DNA and protein sequences used in gene and anti-sense therapy.

PS Claim 24; Page 168-170; 238pp; English.

A method for detecting the presence of neural thread protein (NTP) having a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human subject comprises (a) contacting a sample from a human subject that is suspected of containing the NTP with at least one molecule capable of binding to the protein; and (b) detecting any of the molecule bound to the protein. The binding molecule is selected from an antibody free of natural impurities, a monoclonal antibody or a binding fragment of either of these. The method may be used for diagnosing the presence of Alzheimer's disease, neuroectodermal tumours and a malignant astrocytoma in a human

SQ Sequence 1418 BP; 302 A; 396 C; 315 G; 405 T; 0 U; 0 Other;

### Query Match

Best Local Similarity 96.9%; Pred. No. 0;

Matches 1375; Conservative 0; Mismatches

**Qy** 2 TTTTCTTTTTTGAGATGGGAGTTTTCGCTCTGTGTTGCCAGGCTGGAGTGCAATGGCGCAAT 61

QY 1136 CAGGTGACCCAGCTGCTCAGCTTCCAAAGTGTGGGATTACAGCGTGAGCCACCTCA 1195  
 Db |||||  
 QY 1137 CAGGTGACCCAGCTGCTCAGCTTCCAAAGTGTGGGATTACAGCGTGGA--CGCTCA 1194  
 Db |||||  
 QY 1196 CCAGCGGCTAATTTAGATATAAAATAATGTAGCAATGGGGTCTTCTATGTTGCC 1255  
 Db |||||  
 QY 1195 CCAGCGGCTAATTTAGATATAAAATAATGTAGCAATGGGGTCTTCTATGTTGCC 1254  
 Db |||||  
 QY 1256 AGGCTGGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACCAACCCAGCC 1315  
 Db |||||  
 QY 1255 AGGCTGGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACCAACCCAGCC 1314  
 Db |||||  
 QY 1316 AGTCACATTTTAAACAGTTACATCTTTATTTAGTATACAGCAATGAAATGAAATGAA 1375  
 Db |||||  
 QY 1315 AGTCACA-TTTTAAACAGTTACATCTTTATTTAGTATACAGCAATGAAATGAAATGAA 1373  
 Db |||||  
 QY 1376 CATGTCAAACTCGCAATTCAGTAGTAACAGAGTCTTT 1414  
 Db |||||  
 QY 1374 ATGGCGGAACCTGCAATTCGAGTAGTACAGAGTCTTT 1412  
 Db |||||

RESULT 10

AAQ77883  
 ID AAQ77883 standard; cDNA; 1381 BP.

AC AAQ77883;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 06-JUL-1995 (first entry)  
 XX  
 XX Neural thread protein AD10-7 cDNA.  
 XX  
 XX Neural thread protein AD10-7; Alzheimer's; neuroectodermal tumours;  
 KW malignant astrocytomas; glioblastomas; ss.  
 XX  
 XX Rattus rattus.  
 OS  
 XX  
 XX WO9423756-A1.  
 XX  
 XX 27-OCT-1994.

XX 20-APR-1994; 94WO-US004321.

XX 20-APR-1993; 93US-00050559.

XX (GEO ) GEN HOSPITAL CORP.

XX De La Monte SM, Wanda JR;

XX WPI; 1994-341497/42.

XX Detection of neural thread proteins - to detect sporadic and familial  
 PT Alzheimer's disease, neuroectodermal tumours, malignant astrocytomas and  
 PT glioblastomas (Eng).

PS Example 4; Fig 16R; 158pp; English.

XX AAQ77883 is the AD10-7 neural thread protein (NTP) cDNA. This sequence  
 CC was used in the development of an antibody dependent method, for the  
 CC detection of NTPs. This new method could be used to diagnose Alzheimer's  
 CC disease (differentiating between sporadic and familial), neuroectodermal  
 CC tumours, malignant astrocytomas and glioblastomas. (Updated on 25-MAR-  
 CC 2003 to correct PN field.)

XX Sequence 1381 BP; 296 A; 385 C; 301 G; 399 T; 0 U; 0 Other;

Query Match 74.8%; Score 1078.6; DB 2; Length 1381;

Best Local Similarity 94.3%; Pred. No. 6.3e-279; Mismatches 25; Gaps 18;

Matches 1316; Conservative 0; Indels 54;

QY 2 TTTTATTTTGTAGATGGAGTTTCGCTCTGTTGCGGCTGAGTGCAATGGCGCAAT 61  
 |||||

Db |||||  
 QY 1 TTTTATTTTGTAGATGGAGTTTCGCTCTGTTGCGGCTGAGTGCAATGGCGCAAT 60  
 QY 62 CTCAGCTCACCGCAACCTCCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGCCTCCCC 121  
 Db |||||  
 QY 61 CTCAGCTCACCGCAACCTCCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGCCTCCCC 120  
 QY 122 AGTAGCTGGGATTACAGGCATGTGCACCCAGCTCGGCTAAATTTTGTATTTTATTAGT 181  
 Db |||||  
 QY 121 AGTAGCTGGGATTACAGGCATGTGCA--CCAGCTCGGCTAAATTTTGTATTTTATTAGT 179  
 QY 182 AGAGATGAGATTTT--CTCCATGTTGGTTCAGGCTGCTCGAACTCCGACCTCAGATGAT 239  
 Db |||||  
 QY 180 AGAGATGAGATTTTAACTCAATGTTGGTTCAGGCTGCTCGAACTCCGACCTCAGATGAT 239  
 QY 240 CCTCCGCTCTCGGCTCCCAAAGTCT--AGATACAGAGCTGGGCCACCATGCCCCGG--CT 295  
 Db |||||  
 QY 240 CTCCCGTCTCGGCTGCCCCAAAGTCTCAGATTACAGGCATGAGCCACCATGCCCCGGCT 299  
 QY 296 CTGCTGCTGCTAAATTTTGTGGTGTAGAAACAGGGTTTCACTGATG--TGCCCCAGCTGCTC 354  
 Db |||||  
 QY 300 CTGCTGCTGCTAAATTTTGTGGTGTAGAAACAGGGTTTCACTGATGTTGCCCAAGCTGCTC 359  
 QY 355 CTGAGCTCAAGCAGTCCACCTCGCTCAGCTCCCAAAGTGTGGGATTACAGGCGTGCAG 414  
 Db |||||  
 QY 360 CTCAGCTCAAGCAGTCCACCTCGCTCAGCTCCCAAAGTGTGGGATTACAGGCGT--CAG 418  
 QY 415 CCGTCCCTGGGCTTTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGG 474  
 Db |||||  
 QY 419 CCGTCCCTGGGCTTTTATTTTATTTTAAAGACACAGGTGTACCACTCTTACCCAGG 478  
 QY 475 ATGAAGTCAGTGTGTGATCACAGCTCAGTGCAGCTTCACTCTGAGATCAAGCAAT 533  
 Db |||||  
 QY 479 ATGAAGTCAGTGTGTGATCACAGCTCAGTGCAGCTTCACTCTGAGATCAAGCAAT 538  
 QY 534 CCTCTGCTCAGCTCCCAAAGTGTGGGACCAAAGACATGCACCACTACACCTGGTGA 593  
 Db |||||  
 QY 539 CCTCTGCTCAGCTCCCAAAGTGTGGGACCAAAGACATGCACCACTACACCTGG--TA 597  
 QY 594 ATTTTATTTTATTTTAAATTTTGTGGGACCAAAGTGTGGGACCAAAGTGTGGGACCAAAGTGTGGG 653  
 Db |||||  
 QY 598 ATTTTATTTTATTTTAAATTTTGTGGGACCAAAGTGTGGGACCAAAGTGTGGGACCAAAGTGTGGG 656  
 QY 654 GTGAGTGTGGGCAATCTTGGCTCAGTGCACCACTCTGCTCCCGGTTCAAGTATTCTCC 713  
 Db |||||  
 QY 657 GTGAGTGTGGGCAATCTTGGCTCAGTGCACCACTCTGCTCCCGGTTCAAGTATTCTCC 716  
 QY 714 TGCCCCAGCTCTCTGAGTAGTGGGACCTACAGGCGCCCAACCACTAGCTAAATTTT 773  
 Db |||||  
 QY 717 TGCCCCAGCTCTCTGAGTAGTGGGACCTACAGGCGCCCAACCACTAGCTAAATTTT 776  
 QY 774 GTATTTTGTAGTAGATGGGG--TTCAATGTTGCCAGGTTGATCTTGATCTCTGGACC 832  
 Db |||||  
 QY 777 GTATTTTGTAGTAGATGGGGTTTCAATGTTGCCAGGTTGATCTTGATCTCTTGACC 836  
 QY 833 TTGTGATCTGCTGCTCGGCT--CCCAAAGTGTGGGATTACAGCGCTGAGCCACCAG 891  
 Db |||||  
 QY 837 TTGTGATCTGCTGCTCGGCTACCCAAAGTGTGGGATTACAG--GTGAGTCTCCAC 894  
 QY 892 CCGGCTTATTTTAAATTTTGTGGTGTGGGACCAAAGTGTGGGACCAAAGTGTGGGACCAAAGTGTGGG 951  
 Db |||||  
 QY 895 GCGGCTTATTTTAAATTTTGTGGTGTGGGACCAAAGTGTGGGACCAAAGTGTGGGACCAAAGTGTGGG 954  
 QY 952 GAGTGCATGGCCAAATCTCGGCTCAGTGCACCACTCTGCTCCCGGCTCAGGCAATCT 1011  
 Db |||||  
 QY 955 GAGTGCATGG--CAAATCTCGGCTACTCGCAACCTCTGCTCCCGG--TCAAGCGATCTCT 1012  
 QY 1012 CCGTCTCAGCTCCCAAAGCAGCTGGGATTACGGGACCTGCGCACCAACCCCGCTAAAT 1071  
 Db |||||  
 QY 1013 CCGTCTCAGCTCCCAAAGCAGCTGGGATTACGGG--ACCTGCACCAACACCCCGCTAAAT 1070  
 QY 1072 TTTGTATTTTCAATTAGAGCGGGTTTCAACATATTTGTGAGGCT--GGTCTCAAACTCT 1130  
 Db |||||  
 QY 1071 TTTGTATTTTCAATTAGAGCGGG--TTTACCATAATTTGTGAGGCTGGGCTCTCAAACTCTCT 1128  
 Db |||||





FT	allele	/number= 11
FT		replace(24472,C)
FT		/*tag= ao
FT	allele	replace(24542,A)
FT		/*tag= ap
FT	exon	24628. .24784
FT		/*tag= aq
FT	intron	/number= 12
FT		24785. .29392
FT		/*tag= ar
FT		/number= 12
FT	allele	replace(24900,T)
FT		/*tag= as
FT	exon	29393. .29448
FT		/*tag= at
FT		/number= 13
FT	intron	29449. .29701
FT		/*tag= au
FT		/number= 13
FT	exon	29702. .29859
FT		/*tag= av
FT		/number= 14
FT	intron	29860. .29948
FT		/*tag= aw
FT		/number= 14
FT	exon	29949. .30111
FT		/*tag= ax
FT		/number= 15
FT	allele	replace(31542,T)
FT		/*tag= ay
FT	allele	replace(32344,A,G,C)
FT		/*tag= az
XX		
PN	WO2003095612-A2.	
XX		
PD	20-NOV-2003.	
XX		
PP	05-MAY-2003; 2003WO-US013975.	
XX		
PR	06-MAY-2002; 2002US-0380134P.	
XX		
PA	(APPL-) APPLERA CORP.	
XX		
PI	Abu-Threideh J, Neelam B, Yan C;	
XX		
DR	WPI; 2003-903976/82.	
DR	P-PSDB; ABM79781.	

Qy	64	CAGCTCACCGCAACTCCGCTCCGGGTTCAAGCGATTCTCTGCTCAGCTCCTCCAG	123
Db	32716	CAGCTCACCAAACTTCGCTCCGAGTTCAAGCAAATCTCTCACTCAGCTCCGAG	32657
Qy	124	TAGCTGGGATTACAGGCATGTGCACCCACGCTCGGCTAAATTTTGTATTTTTTTTAGTAG	183
Db	32656	TAGCTGGGATTACAGGCATGTGTACCAACACTCGCTGCTAAATTTTGTATTTTTTT	32601
Qy	184	AGATGGAGTTTCTCATGTGTGTGAGGTGCTCGAACTCCCGACTCAGATGATCCCT	243
Db	32600	AGA CGGGGTTTCTCAATGTGTGTGAGGTGCTTTGTAACCCCGACTCAGATGATCCGC	32541
Qy	244	CCGCTCGGCTCCCAAAGTGTAGATACAGGACTGGCCACCATGCCCGCTCTGCCTGG	303
Db	32540	CCACTTGGCTCCCAAAGTGTCT-GGGATTACAGGTGTAGTTTTTACAAGAAACATTTA	32482
Qy	304	CTAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTCTGAGCTCA	363
Db	32481	AGTATGAAGCAGTACCTCCGGAATATGACGCGCCTCCAAACCCAGCTCTGAAGACTC	32422
Qy	364	AGCAGTCCACCTGCTCAGCTCCCAAAGTGTGGGATTAACAGGCGTGAGCGCTGCTG	423
Db	32421	TTCACTGTCTGGCTGGGCCACCACTCAGCTCCACCTAAACTGGGCTGTGTTCTCTT	32362
Qy	424	GCC-----TTTTTAATTTTATTTTTTTTAAAGACACAGGTGTCCTCACTC	465
Db	32361	CCCCCTCCATCCCTGCTTTTTTTTTTTTTTTTTTTTTTTTGTGAGACAGAGTCTGTGTTG	32302
Qy	466	TTACCCAGGATGAAGTGAGTGGTGTGATCAAGCTCACTGAGCGCTTCAACTCTCTGAGA	525
Db	32301	TTGGCCAGGCTGGAGTGAGTGGTCAAACTATGGCTCACTGAGCGCTTGTATCTCTTGGGC	32242
Qy	526	TC-AAGCATCTCTGCTCAGCTCCCAAGTGTGGGACCAAGACATGCACCACTAC	584
Db	32241	TCAAGGATCTCCCACTCAGCTCTGAGTAGCTGGGACTATAGGTGCAATGCCACAT	32182
Qy	585	ACCTGGGCTAAATTTTATTTTTTATTTTTTGAAGACAGAGTCTCAACTCTGTCAAC	644
Db	32181	GCCCAACTAAATTTTTTTTTTTTTTTTTTTTTTTTGTGAGATGGAGTTTGTCTGTGCTC	32130
Qy	645	CAGCTGGAGTGAGTGCGGCAATCTTGCTCACTGCAACCTCTGCTCCCGGTTCAAG	704
Db	32129	CAGGCTGGGGTGCAATGGTGGCAATCTCGGCTCACTGCAACCTCTGCTCTCTGGGTTCAAG	32070
Qy	705	TTATTTCTCTGCCCGAGCTCTCTGAGTAGCTGGGACTACAGCGCCCCACACGCTAGCT	764
Db	32069	CAATTTCTCTGCTCAACCCCGCAAGTAGCTGGGATTTGAGGCAATGTCATCACACTG	32010
Qy	765	AAATTTTTTGTATTTTTTAGTAGAGATGGGG--TTCAACCATGTTTCCGAGGTGTGATCTTGAT	823
Db	32009	GCTAAATTTGTATTTGTATTAGTAGAGATGGGTTTCACTATGTCAGTCAGGCTGTCTCAAA	31950
Qy	824	CTCTGGAGCT--TGATGATCTGCTCGGCTCCCAAAGTGTCTGGGATTTACAGGCGTG	881
Db	31949	CTCTTGACTCAGGAGATCTCTCTGCTCGGCTCCCGAGGGGCTGGGATTTACAGGTGTG	31890
Qy	882	AGCCACCAAGCCCGGCTT---ATTTTTAAATTTTGTGTTTGAATGGAATCTCACTCT	938
Db	31889	AGCCACTGTGCTGGCTGCCCCAATTTTTTTTTTTTTTTTTTTTGTGAGACAAGTCTCGCT	31830
Qy	939	GTTTACCCAGGTGGAGTGCAATGGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGGG	998
Db	31829	CTTTCCAGGCTGGAGTGGATGGCAACCACTCGGCTCACTGCAACCTCTGCTTTGCGG	31770
Qy	999	CTCAAGCGATTTCTCTGTCTCAGCTCCCAAGCAGCTGGGATTTACGGGCACTGCCACCA	1058
Db	31769	TTCAAGCGATTTCTCTGCTCAGTCTCCCAAGTAGTAGTACAGGCAATGTGCCACCA	31710
Qy	1059	CACCCCGCTAATTTTGTATTTTCAATTAGCGGGGTTTTCACCATATTTGTCAAGCTGG	1118
Db	31709	CGCCCGGCTAATTTTGTATTTTATAGTAGACGAGTTTTTCAACATGTTGGCGCAGACTGG	31650
Qy	1119	TCTCAAACTCTCTGACCTCAGGTGACCACTCGCTCAGCTTTTCCAAAGTCTGGGATTC	1178

Db 31649 TTTGACTCTTAATCTAGGTGATCCACCGCCTCTGCTCCCAAGTCTGGGATTAC 31590  
 QY 1179 AGCGTGAGCCACCTCACCCAGCC-----GGCTAATTTAGATAAAAAAATATGTAGCAATG 1234  
 Db 31589 AGGCATGAGCCACACACCTGGCCTCTGCCCAACTAATTTAAAAAATTTTGTAGAGA 31530  
 QY 1235 GGGGCTCTGCTATGTTGCCAGGCTGGTCTCAAACTTCTGCTTCAATGCAATCCTTCA 1294  
 Db 31529 CAGGATCTCACTATGTTGTCAGACTGGTCTTGAACCTCCTGGGCTTAAGGGATCGTCCTG 31470  
 QY 1295 AATGAGCCAC 1304  
 Db 31469 TCTCAGCCTC 31460

RESULT 13  
 ABL62910/c  
 ID ABL62910 standard; DNA; 65608 BP.  
 XX  
 AC ABL62910;  
 XX  
 DT 15-MAY-2002 (first entry)  
 XX  
 DE Breast cancer related gene sequence SEQ ID NO:1247.  
 XX  
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
 KW gene; db.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200194629-A2.  
 XX  
 PD 13-DEC-2001.  
 XX  
 PF 30-MAY-2001; 2001WO-US010838.  
 XX

PR 05-JUN-2000; 2000US-0209473P.  
 PR 05-JUN-2000; 2000US-0209531P.  
 PR 18-SEP-2000; 2000US-0233131P.  
 PR 18-SEP-2000; 2000US-0233617P.  
 PR 20-SEP-2000; 2000US-0234009P.  
 PR 20-SEP-2000; 2000US-0234034P.  
 PR 20-SEP-2000; 2000US-0234052P.  
 PR 22-SEP-2000; 2000US-0234509P.  
 PR 22-SEP-2000; 2000US-0234567P.  
 PR 25-SEP-2000; 2000US-0234923P.  
 PR 25-SEP-2000; 2000US-0234924P.  
 PR 25-SEP-2000; 2000US-0235077P.  
 PR 25-SEP-2000; 2000US-0235082P.  
 PR 25-SEP-2000; 2000US-0235134P.  
 PR 25-SEP-2000; 2000US-0235280P.  
 PR 26-SEP-2000; 2000US-0235637P.  
 PR 26-SEP-2000; 2000US-0235638P.  
 PR 27-SEP-2000; 2000US-0235711P.  
 PR 27-SEP-2000; 2000US-0235720P.  
 PR 27-SEP-2000; 2000US-0235840P.  
 PR 27-SEP-2000; 2000US-0235863P.  
 PR 28-SEP-2000; 2000US-0236028P.  
 PR 28-SEP-2000; 2000US-0236032P.  
 PR 28-SEP-2000; 2000US-0236033P.  
 PR 28-SEP-2000; 2000US-0236034P.  
 PR 28-SEP-2000; 2000US-0236109P.  
 PR 28-SEP-2000; 2000US-0236111P.  
 PR 29-SEP-2000; 2000US-0236842P.  
 PR 29-SEP-2000; 2000US-0236891P.  
 PR 02-OCT-2000; 2000US-0237172P.  
 PR 02-OCT-2000; 2000US-0237278P.  
 PR 02-OCT-2000; 2000US-0237294P.  
 PR 02-OCT-2000; 2000US-0237295P.

PR 02-OCT-2000; 2000US-0237316P.  
 PR 03-OCT-2000; 2000US-0237425P.  
 PR 03-OCT-2000; 2000US-0237598P.  
 PR 03-OCT-2000; 2000US-0237604P.  
 PR 03-OCT-2000; 2000US-0237606P.  
 PR 03-OCT-2000; 2000US-0237608P.  
 PR 01-NOV-2000; 2000US-0244867P.  
 PR 01-NOV-2000; 2000US-0245084P.  
 XX  
 PA (AVAL-) AVALON PHARM.  
 XX  
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Soppet DR, Weaver Z;  
 XX  
 DR WPI; 2002-188264/24.  
 XX  
 PT Screening for anti-neoplastic agent involves exposing cells to a chemical  
 PT agent to be tested for anti-neoplastic activity, and determining a change  
 PT in expression of a gene of a signature gene set.  
 XX  
 PS Claim 1; SEQ ID NO 1247; 44pp; English.  
 XX  
 CC The present invention describes a method (M1) for screening for an anti-  
 CC neoplastic agent. The method involves exposing cells to a chemical agent  
 CC to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening an  
 CC anti-neoplastic agent, and can be used for producing a product which is  
 CC the data collected with respect to the anti-neoplastic agent as a result  
 CC of M1, and the data is sufficient to convey the chemical structure and/or  
 CC properties of the agent. M1 can be used in the treatment of cancer such  
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's  
 CC tumour  
 XX  
 SQ Sequence 65608 BP; 15987 A; 16979 C; 16761 G; 15878 T; 0 U; 3 Other;  
 Query Match 40.8%; Score 587.8; DB 6; Length 65608;  
 Best Local Similarity 70.6%; Pred. No. 1.6e-146;  
 Matches 940; Conservative 0; Mismatches 347; Indels 44; Gaps 10;  
 QY 4 TTTTGTGAGATGGAGTTTTCGCTCTTGTGTCGCCAGGCTGAGTGCAATGCGCAATCT 63  
 Db TTTTGTGAGATGGAGTTTTCGCTCTTGTGTCGCCAGGCTGAGTGCAATGCGCAATCT 39935  
 QY 64 CAGCTCACCGCAACCTCCGCTCCCGGTTTCAAGCGATTCTCTGCTCAGCTCCCGAG 123  
 Db CAGCTCACCGCAACCTCCGCTCCCGGTTTCAAGCGATTCTCTCCTACCTCAGCTCCCGAG 39875  
 QY 124 TAGCTGGGATTTACAGGATGTGCACCCAGGCTCGGCTAATTTTGTATTTTGTAGTAG 183  
 Db TAGCTGGGATTTACAGGATGTGTCCACCACTGGCTAATTTTGTATTTTGTATTTT 39819  
 QY 184 AGATGGAGTTTCTCCATGTTGTCAGGCTGTCTCGAACTCCGACCTCAGATGATCCT 243  
 Db AGACGGGTTTCTCCATGTTGTCAGGTTGTCTTGAACCCCGACCTCAGATGATCCTG 39759  
 QY 244 CCGTCTCGGCTCCCAAGTCTAGATACAGGACTGGCCACATGCCCGGCTCTGCCCTGG 303  
 Db CCACCTTGGCTCCCAAGTCT-GGGATTACAGGTGTGAGTTTTCACAGAAACATTTA 39700  
 QY 304 CTAATTTTGTGTAGAAACAGGGTTTTCATGATGTGCCAGCTGTGTTCTCTGAGCTCA 363  
 Db AGTATGAAAGCAGTACCTCCGGAATATGACGGGCTCCCAACCCAGCTCTGAAAGCTC 39640  
 QY 364 AGCAGTCCACCTGCTCAGCTCCCAAGTGTCTGGATTA----- 403  
 Db TTCACCTTGTGCTGGCTGGGCCACCACTCAGCTCCACCTAAACTGGGCTGTGTCTCTT 39580





CC neoplastic agent. The method involves exposing cells to a chemical agent  
CC to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 847 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. MI can be used for screening an  
CC anti-neoplastic agent, and can be used for producing a product which is  
CC the data collected with respect to the anti-neoplastic agent as a result  
CC of MI, and the data is sufficient to convey the chemical structure and/or  
CC properties of the agent. MI can be used in the treatment of cancer such  
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's  
CC tumour  
XX  
SQ Sequence 65608 BP; 15987 A; 16979 C; 16761 G; 15878 T; 0 U; 3 Other;  
Query Match 40.8%; Score 587.8; DB 6; Length 65608;  
Best Local Similarity 70.6%; Pred. No. 1.6e-146;  
Matches 940; Conservative 0; Mismatches 347; Indels 44; Gaps 10;  
QY 4 TTTTGTGAGATGGAGTTTTCGCTCTTGTGCTCCAGGCTGAGTGCATGGCGCAATCT 63  
DB 39993 TTTTGTGAGATGGAGTTT-GCTCTGTGCTCCAGGCTGAAGTGAATGGCAGATCT 39935  
QY 64 CAGCTCACCGCAACTCCGCCCTCCGGGTTCAAGCGATTCTCCTGCCTCAGCTCCCGAG 123  
DB 39934 CAGCTCACCGCAACTCCGCCCTCCAGGTTCAAGCAATCTCTACCTCAGCTCCCGAG 39875  
QY 124 TAGCTGGGATTACGGCATGTGCAACCGCTCGCTAATTTTGTATTTTGTAGTAG 183  
DB 39874 TAGCTGGGATTACGGCATGTGTCACCACTGGCTAATTTTGTATTTT-----AGT 39819  
QY 184 AGATGAGATTCTCCATGTGCTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATCCCT 243  
DB 39818 AGACGGGTTTCTCCATGTGCTCAGGTTGGTCTTGAACCCCGACCTCAGATGATCCG 39759  
QY 244 CCGTCTCGGCCCTCCAAAGTGTAGATACAGGACTGGCCACCATGCCGGCTCTCGCTGG 303  
DB 39758 CCACCTTGGCCCTCCCAAGTGTG-GGGATTACAGGTGTGAGTTTACAGAAACATTTA 39700  
QY 304 CTAAATTTTGTGGTAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTCGAGCTCA 363  
DB 39699 AGTATGAACAGATACCTCCGGAATATGCAGGCCCTCCAAACCCAGCCTCTGAAGACTC 39640  
QY 364 AGCAGTCCACCTCGCTCAGCCTCCCAAGTGTGGGATTA----- 403  
DB 39639 TTCACTTGTCTGGCTGGCCACCACTCAGCTCCACCTTAAACTGGGCTGTCTCTTT 39580  
QY 404 CAGGGGTGACGGCTGGCTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCAC 463  
DB 39579 CCCCCCTCCATCCCTGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 39520  
QY 464 TCTTACCCAGGATGAGTGCAGTGGTGTGATCAGAGTCACTGCAGCCTTCACTCCCTGA 523  
DB 39519 TGTGTGCCAGGTGGAGTGCAGTGGTGCAGAAATGGCTCACTGCAGCCTTGTATCTTGG 39460  
QY 524 GATC-AAGCATCTCTCGCTCAGCCTCCCAAGTGTGGGACCAAGACATGCACCACT 582  
DB 39459 GCTCAAGGATCTCCCACTCAGCTCCTGAGTGTAGTGGGATATAGTGCATGCCACC 39400  
QY 583 ACACCTGGCTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 642  
DB 39399 ATGCCCAACTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 39348  
QY 643 CCGAGCTGGAGTGCAGTGGCGCAATCTTGGCTCACTGCAGCACTCGCTCCCGGGTTCA 702  
DB 39347 CCGAGCTGGGTTGCAATGGTGGATCTCGGCTCACTGCAGCACTCGCTCCCGGGTTCA 39288  
QY 703 AGTTATTTCTCTCGCCAGCTCCTCAGTGTAGTGGGACTACAGGCGCCCAACCAAGCTAG 762

DB 39287 AGCAATTTCTCTGCTCAACCCCAAGTAGCTGGGATTTTCAGGCATGTGCCATCACACC 39228  
QY 763 CTAAATTTTGTATTTTATTTAGTAGAGATGGG-TTCACCATGTTGCGCAGGTGATCTTG 821  
DB 39227 TGGCTAAATTTTGTATTTAGTAGAGATGGGTTTCACTATGTAGTCAAGCTGTCTCA 39168  
QY 822 ATCTCTGACCT--TGTGATCTGCTGCTCGGCTCCCAAGTGTGGGATTTACAGCG 879  
DB 39167 AACTCTGACCTCAGGATCTCTGCTCGGCTCCCAAGTGTGGGATTTACAGGTG 39108  
QY 880 TGAGCCACCGCCGGCTT--ATTTTAAATTTTGTGTTTGAATGGAATCTCACTC 937  
DB 39107 TGAGCCACTGTGCTGGCTGCCAATTTTATTTTATTTTATTTTATTTTATTTTATTT 39048  
QY 938 TGTACCAGCTGGAGTGCATGGCCAAATCTGGCTCACTGCACCTCTGCTCCCGG 997  
DB 39047 TCTTTCCAGGCTGGAGTGGAAATGGCACCTCGGCTCACTGCACCTCTGCTTTGG 38988  
QY 998 GCTCAAGGATTTCTCTGCTCAGCTCCCAAGCAGCTGGGATTTACGGGCACTGCCACC 1057  
DB 38987 GTTCAAGGATTTCTCTGCTCAGCTCCCAAGTAGCTAGGACTACAGGCACTGCCACC 38928  
QY 1058 ACACCCGCTAATTTTGTATTTTATTTAGAGGGGGTTCACATATTTGTCAAGGTG 1117  
DB 38927 AGCCCGGCTAATTTTGTATTTTATTTAGTAGAGACGAGTTTTCACCATGTGGCCAG 38868  
QY 1118 GTCTCAACTCTGACCTCAGTGCACCTGCTCAGCTTCCAAAGTGTGGGATTA 1177  
DB 38867 GTCTTGAACCTCTAATTTCAAGGTGATCCACCGCTCTGCTCCCAAGTGTGGGATTA 38808  
QY 1178 CAGGCTGAGCCACTCACCACCGC-----GGCTAATTTAGATAAAAAATATGTAGCAAT 1233  
DB 38807 CAGGATGAGCCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 38748  
QY 1234 GGGGGTCTTGTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1293  
DB 38747 ACAGGATCTCACTATTTGTCCAGACTGCTTGAACCTCTGCTGCTGCTGCTGCTGCT 38688  
QY 1294 AATGAGCCAC 1304  
DB 38687 GTCTCAGCCTC 38677  
RESULT 15  
ABL67668/c  
ID ABL67668 standard; DNA; 65608 BP.  
XX ABL67668;  
AC ABL67668;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Oesophagus cancer related gene sequence SEQ ID NO:6005.  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX Homo sapiens.  
OS  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US010838.  
XX  
PR 05-JUN-2000; 2000US-0209473P.  
PR 05-JUN-2000; 2000US-0209531P.  
PR 18-SEP-2000; 2000US-023133P.  
PR 18-SEP-2000; 2000US-0233617P.  
PR 20-SEP-2000; 2000US-0234009P.  
PR 20-SEP-2000; 2000US-0234034P.  
PR 20-SEP-2000; 2000US-0234052P.





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 16:58:08 ; Search time 1049 Seconds  
(without alignments)  
9158.654 Million cell updates/sec

Title: US-09-380-203-1  
Perfect score: 1442  
Sequence: 1 TTTTCTTTTTCGATGCGAG.....TTAAACAAAGCTTAGACGA 1442

Scoring table: IDENTITY\_NUC  
Gap 10.0, Gapext 1.0

Searched: 7389322 seqs, 3331285599 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10J\_PUBCOMB.seq:\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1442	100.0	1442	9	US-09-964-666-1 Sequence 1, Appli
2	1442	100.0	1442	9	US-09-964-412-1 Sequence 1, Appli
3	1442	100.0	1442	10	US-09-964-667-1 Sequence 1, Appli
4	1442	100.0	1442	10	US-09-872-968-1 Sequence 1, Appli
5	1442	100.0	1442	10	US-09-964-678A-1 Sequence 1, Appli
6	1442	100.0	1442	14	US-10-146-130-1 Sequence 1, Appli
7	1442	100.0	1442	14	US-10-092-934-1 Sequence 1, Appli

8	1442	100.0	1442	14	US-10-153-334-53 Sequence 53, Appli
9	1442	100.0	1442	14	US-10-198-069-48 Sequence 48, Appli
10	1442	100.0	1442	15	US-10-198-070-125 Sequence 125, App
11	1442	100.0	1442	19	US-10-755-889-409 Sequence 409, App
12	1442	100.0	1442	21	US-10-910-173-1 Sequence 1, Appli
13	1223.4	84.8	1418	9	US-09-964-666-4 Sequence 4, Appli
14	1223.4	84.8	1418	9	US-09-964-412-4 Sequence 4, Appli
15	1223.4	84.8	1418	10	US-09-964-667-4 Sequence 4, Appli
16	1223.4	84.8	1418	10	US-09-964-678A-4 Sequence 4, Appli
17	1080.2	74.9	1381	9	US-09-964-666-3 Sequence 3, Appli
18	1080.2	74.9	1381	9	US-09-964-412-3 Sequence 3, Appli
19	1080.2	74.9	1381	10	US-09-964-667-3 Sequence 3, Appli
20	1080.2	74.9	1381	10	US-09-964-678A-3 Sequence 3, Appli
21	590.4	40.9	33112	17	US-10-429-873A-3 Sequence 3, Appli
22	587.8	40.8	65608	9	US-09-962-436-292 Sequence 292, App
23	587.8	40.8	65608	9	US-09-962-832-119 Sequence 119, App
24	587.8	40.8	65608	9	US-09-954-531-180 Sequence 180, App
25	587.8	40.8	65608	21	US-10-843-641A-1247 Sequence 1247, App
26	587.8	40.8	65608	21	US-10-843-641A-2751 Sequence 2751, Ap
27	587.8	40.8	65608	21	US-10-843-641A-6005 Sequence 6005, Ap
28	535.2	37.1	79528	20	US-10-723-860-2621 Sequence 2621, Ap
29	535.2	37.1	79528	22	US-10-756-149-2427 Sequence 2427, Ap
30	524.4	36.4	62944	9	US-09-954-456-2257 Sequence 2257, Ap
31	524.4	36.4	62944	11	US-09-968-007A-129 Sequence 129, App
32	524.4	36.4	62944	21	US-10-843-641A-5284 Sequence 5284, Ap
33	524.4	36.4	62944	21	US-10-843-641A-6599 Sequence 6599, Ap
34	513.4	35.6	28805	13	US-10-087-192-2008 Sequence 2008, Ap
35	513	35.6	21470	9	US-09-764-847-1157 Sequence 1157, Ap
36	513	35.6	21470	14	US-10-092-154-1157 Sequence 1157, Ap
37	502.8	34.9	75252	13	US-10-087-192-304 Sequence 904, App
38	501	34.7	193691	20	US-10-719-993-6768 Sequence 6768, Ap
39	496.4	34.4	115756	22	US-10-756-149-3838 Sequence 3838, Ap
40	490.8	34.0	166496	22	US-10-756-149-590 Sequence 590, App
41	490.4	34.0	99014	9	US-09-880-107-3428 Sequence 3428, Ap
42	490.2	34.0	130001	17	US-10-199-676-23 Sequence 23, Appli
43	490.2	34.0	130001	21	US-10-639-300-23 Sequence 23, Appli
44	490	34.0	51837	19	US-10-322-281-598 Sequence 598, App
45	490	34.0	76201	17	US-10-439-703-84 Sequence 84, Appli

## ALIGNMENTS

RESULT 1  
US-09-964-666-1  
; Sequence 1, Application US/09964666  
; Patent No. US20020104108A1  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; Wands, Jack R.  
; TITLE OF INVENTION: Transgenic Animals and Cell Lines for  
; Screening Drugs Effective for the Treatment or Prevention  
; of Alzheimer's Disease  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.  
; STREET: 1100 New York Ave., Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/09/964,666  
; FILING DATE: 28-Sep-2001  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0609.4370000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1442 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: both  
MOLECULE TYPE: cdna  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 15..1139  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-964-666-1

Query Match 100.0%; Score 1442; DB 9; Length 1442;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTT...TTTGGAGTGGAGTTTGGCTCTTGTGGCCGAGGCTGGAGTGCATATGGCGCA 60  
Db 1 TTTT...TTTGGAGTGGAGTTTGGCTCTTGTGGCCGAGGCTGGAGTGCATATGGCGCA 60

QY 61 TCTCAGCTCAGCGCAACCTCCGCTCCCGGGTTCAAGCGATTCTCCTGCTCAGCTCC 120  
Db 61 TCTCAGCTCAGCGCAACCTCCGCTCCCGGGTTCAAGCGATTCTCCTGCTCAGCTCC 120

QY 121 CAGTAGTGGGATTACAGCATGTGCACCCAGCTCGGCTAAATTTTGTATTTTTTTAG 180  
Db 121 CAGTAGTGGGATTACAGCATGTGCACCCAGCTCGGCTAAATTTTGTATTTTTTTAG 180

QY 181 TAGAGTGGAGTTTCTCATGTTGTTCAGGCTGGTCTGAACTCCGACCTCGATGATC 240  
Db 181 TAGAGTGGAGTTTCTCATGTTGTTCAGGCTGGTCTGAACTCCGACCTCGATGATC 240

QY 241 CTTCCGCTCTCGGCTCCCAAGTGTCTAGATACAGACTGGCCACCATGCCGGCTCGC 300  
Db 241 CTTCCGCTCTCGGCTCCCAAGTGTCTAGATACAGACTGGCCACCATGCCGGCTCGC 300

QY 301 TGGCTAAATTTTGTGGTGAACAAGGTTTCACTGATGTGCCCAAGCTGCTCTGAGC 360  
Db 301 TGGCTAAATTTTGTGGTGAACAAGGTTTCACTGATGTGCCCAAGCTGCTCTGAGC 360

QY 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGGGATTACAGGCGTGGAGCGTGC 420  
Db 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGGGATTACAGGCGTGGAGCGTGC 420

QY 421 CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480  
Db 421 CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480

QY 481 TGCAGTGTGTGATCAGAGTCACTGCAGCTTCAACTCTGAGATCAAGCATCTCTG 540  
Db 481 TGCAGTGTGTGATCAGAGTCACTGCAGCTTCAACTCTGAGATCAAGCATCTCTG 540

QY 541 CTTGAGCTTCCCAAGTGTGGGACCAAGATGACCACTACACCTGGCTAAATTTTA 600  
Db 541 CTTGAGCTTCCCAAGTGTGGGACCAAGATGACCACTACACCTGGCTAAATTTTA 600

QY 601 TTTT...TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660  
Db 601 TTTT...TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660

QY 661 GCGCAATCTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTATTTCTCTGCCCA 720  
Db 661 GCGCAATCTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTATTTCTCTGCCCA 720

QY 721 GCCTCTGAGTGTGGGACTACAGCGGCCACAGCGCTAGCTAATTTTGTATTTT 780  
Db 721 GCCTCTGAGTGTGGGACTACAGCGGCCACAGCGCTAGCTAATTTTGTATTTT 780

781 TAGTAGAGATGGGTTTCCACATGTTTCGACAGTGTGATCTTGATCTCTGACCTTGTGATC 840  
Db 781 TAGTAGAGATGGGTTTCCACATGTTTCGACAGTGTGATCTTGATCTCTGACCTTGTGATC 840

QY 841 TGCCTGCTCGGCTTCCCAAGTGTGGGATTACAGGCGTGGAGCCACCAAGCGGCTTA 900  
Db 841 TGCCTGCTCGGCTTCCCAAGTGTGGGATTACAGGCGTGGAGCCACCAAGCGGCTTA 900

QY 901 TTTT...TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 960  
Db 901 TTTT...TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 960

QY 961 GGCCTGCTCGGCTCACTGCAACCTCTGCTCCCGGGTCAAGCGATTCTCTCTCTCA 1020  
Db 961 GGCCTGCTCGGCTCACTGCAACCTCTGCTCCCGGGTCAAGCGATTCTCTCTCTCA 1020

QY 1021 GCCTCCCAAGCAGCTGGGATTACAGGCGCTGCGGCTGCGGCTGCGGCTAATTTTGTATTT 1080  
Db 1021 GCCTCCCAAGCAGCTGGGATTACAGGCGCTGCGGCTGCGGCTAATTTTGTATTT 1080

QY 1081 TCATTAGAGGCGGGTTTTCACATATTTGTGAGGCTGCTCAAACTCTGACCTCAGGT 1140  
Db 1081 TCATTAGAGGCGGGTTTTCACATATTTGTGAGGCTGCTCAAACTCTGACCTCAGGT 1140

QY 1141 GACCCACCTGCTCAGGCTTCCAAAGTGTGGGATTACAGGCGTGGAGCCACCTCACCCAG 1200  
Db 1141 GACCCACCTGCTCAGGCTTCCAAAGTGTGGGATTACAGGCGTGGAGCCACCTCACCCAG 1200

QY 1201 CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGTCTTGTCTATGTTGCCAGGCT 1260  
Db 1201 CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGTCTTGTCTATGTTGCCAGGCT 1260

QY 1261 GGTCTCAAACTTGGCTTTCATGCAATCTTCCAAATGAGCCACCAAGCGGCTCA 1320  
Db 1261 GGTCTCAAACTTGGCTTTCATGCAATCTTCCAAATGAGCCACCAAGCGGCTCA 1320

QY 1321 CATTTT...TAAACAGTTACATCTTTATTTAGTACTAGAAAGTAAATAACAATTAACATGT 1380  
Db 1321 CATTTT...TAAACAGTTACATCTTTATTTAGTACTAGAAAGTAAATAACAATTAACATGT 1380

QY 1381 CAAACCTGCAAAATTCAGTAGTAACAGAGTCTTTTATACTTTTAAACAAAGCTTTAGAG 1440  
Db 1381 CAAACCTGCAAAATTCAGTAGTAACAGAGTCTTTTATACTTTTAAACAAAGCTTTAGAG 1440

QY 1441 CA 1442  
Db 1441 CA 1442

RESULT 2  
US-09-964-412-1  
; Sequence 1, Application US/09964412  
; Patent No. US2002012931A1  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; Wanda, Jack R.  
; TITLE OF INVENTION: Transgenic Animals and Cell Lines for  
; Screening Drugs Effective for the Treatment or Prevention  
; of Alzheimer's Disease  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.  
; STREET: 1100 New York Ave., Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/964,412  
 FILING DATE: 28-Sep-2001  
 CLASSIFICATION: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Emond, Robert W.  
 REGISTRATION NUMBER: 32,893  
 REFERENCE/DOCKET NUMBER: 0609,4370000  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-371-2600  
 TELEFAX: 202-371-2540  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1442 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: both  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 15..1139  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-964-412-1

Query Match 100.0%; Score 1442; DB 9; Length 1442;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTCTTTTTCAGATGGAGTTTCGCTCTGTGTGCCAGGCTGGAGTGCATATGGCGAA 60  
 DB 1 TTTTCTTTTTCAGATGGAGTTTCGCTCTGTGTGCCAGGCTGGAGTGCATATGGCGAA 60

QY 61 TCTCAGCTCAGCGCAACCTCCGCTCCCGGGTTCAAGCGATTCTCTGTGCTCAGCGTCC 120  
 DB 61 TCTCAGCTCAGCGCAACCTCCGCTCCCGGGTTCAAGCGATTCTCTGTGCTCAGCGTCC 120

QY 121 CAGTAGTGGGATTCAGGATGTGCAACCGCTCGGCTAAATTTGTTATTTTTTTTAG 180  
 DB 121 CAGTAGTGGGATTCAGGATGTGCAACCGCTCGGCTAAATTTGTTATTTTTTTTAG 180

QY 181 TAGAGATGAGTTTCTCCATGTTGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC 240  
 DB 181 TAGAGATGAGTTTCTCCATGTTGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC 240

QY 241 CTTCTGCTCTCGGCTCCCAAGTGTAGATACAGGACTGGCCACCACTGCGGCTCTGCC 300  
 DB 241 CTTCTGCTCTCGGCTCCCAAGTGTAGATACAGGACTGGCCACCACTGCGGCTCTGCC 300

QY 301 TGGCTAAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTGAGC 360  
 DB 301 TGGCTAAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTGAGC 360

QY 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGGGATTCAGGCGTGCAGCGTGC 420  
 DB 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGGGATTCAGGCGTGCAGCGTGC 420

QY 421 CTGGCCCTTTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGA 480  
 DB 421 CTGGCCCTTTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGA 480

QY 481 TGCAGTGGTGTATCAGCTCACTGACGCTTCACTCTGAGATCAAGCATCTCTCTG 540  
 DB 481 TGCAGTGGTGTATCAGCTCACTGACGCTTCACTCTGAGATCAAGCATCTCTCTG 540

QY 541 CCTCAGCTCCCAAGTGTGGGACCAAGATGCACCACTACACCTGGCTAAATTTT 600  
 DB 541 CCTCAGCTCCCAAGTGTGGGACCAAGATGCACCACTACACCTGGCTAAATTTT 600

QY 601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660  
 DB 601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660

QY 661 GCGCGAATCTCGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCA 720  
 DB 661 GCGCGAATCTCGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCA 720

DB 661 GCGCGAATCTCGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCA 720  
 QY 721 GCCTCTGAGTAGCTGGGACTACAGCGCCACACAGCTAGCTAATTTTTTTTATTTT 780  
 DB 721 GCCTCTGAGTAGCTGGGACTACAGCGCCACACAGCTAGCTAATTTTTTTTATTTT 780  
 QY 781 TAGTAGAGATGGGGTTCAACATGTTGCGCAGGTTGATCTTGATCTCTGGAACCTTGGATC 840  
 DB 781 TAGTAGAGATGGGGTTCAACATGTTGCGCAGGTTGATCTTGATCTCTGGAACCTTGGATC 840

QY 841 TGCCTGCTCCGCTCCCAAGTGTGGGATTCAGGCGTGCAGCCACCGCCGGCTTA 900  
 DB 841 TGCCTGCTCCGCTCCCAAGTGTGGGATTCAGGCGTGCAGCCACCGCCGGCTTA 900

QY 901 TTTTAAATTTTGTGTTGTTGAAATGGGAATCTCACTCTGTTACCCAGGCTGGAGTCAAT 960  
 DB 901 TTTTAAATTTTGTGTTGTTGAAATGGGAATCTCACTCTGTTACCCAGGCTGGAGTCAAT 960

QY 961 GCGCAAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTCTCTGCTCA 1020  
 DB 961 GCGCAAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTCTCTGCTCA 1020

QY 1021 GCCTCCCAAGCAGCTGGGATTCAGGCGCACCTGCGCCACACCGCCGCTAATTTTGTATTT 1080  
 DB 1021 GCCTCCCAAGCAGCTGGGATTCAGGCGCACCTGCGCCACACCGCCGCTAATTTTGTATTT 1080

QY 1081 TCATTAGAGGGGGGTTTCCACATATTTGTTCAGGCTGGTCTCAAACTCTCTGACCTCAGGT 1140  
 DB 1081 TCATTAGAGGGGGGTTTCCACATATTTGTTCAGGCTGGTCTCAAACTCTCTGACCTCAGGT 1140

QY 1141 GACCACCTGCTCAGCTTCCAAAGTGTGGGATTCAGGCGTGCAGCCACCTCAGCCAG 1200  
 DB 1141 GACCACCTGCTCAGCTTCCAAAGTGTGGGATTCAGGCGTGCAGCCACCTCAGCCAG 1200

QY 1201 CCGGCTAATTTAGATAAAAAATATGTAGCAATGGGGGCTCTTGTATGTGCGCCAGGCT 1260  
 DB 1201 CCGGCTAATTTAGATAAAAAATATGTAGCAATGGGGGCTCTTGTATGTGCGCCAGGCT 1260

QY 1261 GGTCTCAAACTTCTGGCTTTCATGCAATCTTCCAAATGAGCCACCAACCCAGCCAGTCA 1320  
 DB 1261 GGTCTCAAACTTCTGGCTTTCATGCAATCTTCCAAATGAGCCACCAACCCAGCCAGTCA 1320

QY 1321 CATTTTTTAAACAGTTACATCTTTATTTTAGTATATAGAAAGTAAATACAATAAATCT 1380  
 DB 1321 CATTTTTTAAACAGTTACATCTTTATTTTAGTATATAGAAAGTAAATACAATAAATCT 1380

QY 1381 CAAACTGCAATTCAGTAGTAAAGAGTTCTTTTATTAATCTTTTAAACAAAGCTTTAGAG 1440  
 DB 1381 CAAACTGCAATTCAGTAGTAAAGAGTTCTTTTATTAATCTTTTAAACAAAGCTTTAGAG 1440

QY 1441 CA 1442  
 DB 1441 CA 1442

RESULT 3

US-09-964-667-1  
 ; Sequence 1, Application US/09964667  
 ; Publication No. US2003003621A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: de la Monte, Suzanne  
 ; Wands, Jack R.  
 ; TITLE OF INVENTION: Transgenic Animals and Cell Lines for  
 ; Screening Drugs Effective for the Treatment or Prevention  
 ; of Alzheimer's Disease  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.  
 ; STREET: 1100 New York Ave., Suite 600  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20005-3934



COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICANT: US/09/964,667  
FILING DATE: 28-Sep-2001  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Emond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0609,4370000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1442 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 15..1139

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-964-667-1

Query Match 100.0%; Score 1442; DB 10; Length 1442;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TTTTATTTTAAATTTTGGAGCAGAGTCTCAACTCTGTCCACCGAGCTGGAGTGCAGT 660

1 TTTTATTTTAAATTTTGGAGCAGAGTCTCAACTCTGTCCACCGAGCTGGAGTGCAGT 660

61 TCTCAGCTCAGCCGACCTCCGCTCCCGGGTTCAAGGATTTCTCTGCTCAGCTCCC 120

61 TCTCAGCTCAGCCGACCTCCGCTCCCGGGTTCAAGGATTTCTCTGCTCAGCTCCC 120

121 CAGTAGTGGGATTACAGGATGTGCACCCAGCTCGGCTAATTTTGTATTTTGTAG 180

121 CAGTAGTGGGATTACAGGATGTGCACCCAGCTCGGCTAATTTTGTATTTTGTAG 180

181 TAGAGATGGAGTTTCTCCATGTGTGTGCTGAGTGTGCTGAACTCCCGACCTCAGATGATC 240

181 TAGAGATGGAGTTTCTCCATGTGTGTGCTGAGTGTGCTGAACTCCCGACCTCAGATGATC 240

241 CCTCGCTCTCGGCTCCCAAGTGTGTAGATAGGACTGGCCACCATGCCGGCTCTGCC 300

241 CCTCGCTCTCGGCTCCCAAGTGTGTAGATAGGACTGGCCACCATGCCGGCTCTGCC 300

301 TGGCTAAATTTTGTGTGTAACACAGGTTTCACTGATGTGCCCAAGCTGGTCTCTCTGAGC 360

301 TGGCTAAATTTTGTGTGTAACACAGGTTTCACTGATGTGCCCAAGCTGGTCTCTCTGAGC 360

361 TCAGGAGTCCAGCTGCTCCCAAGTGTGTAGATAGGACTGGCCACCATGCCGGCTCTGCC 420

361 TCAGGAGTCCAGCTGCTCCCAAGTGTGTAGATAGGACTGGCCACCATGCCGGCTCTGCC 420

421 CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480

421 CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480

481 TGCAGTGTGTGATCAGAGTCTCACTGAGCTTTCAACTCTGAGATCAAGATCTCTCTG 540

481 TGCAGTGTGTGATCAGAGTCTCACTGAGCTTTCAACTCTGAGATCAAGATCTCTCTG 540

541 CTTAGCTTCCCAAGTGTGGGACCAAGAGATGACCACTACACCTTGGCTTATTTTATTTT 600

541 CTTAGCTTCCCAAGTGTGGGACCAAGAGATGACCACTACACCTTGGCTTATTTTATTTT 600

601 TTTTATTTTAAATTTTGGAGCAGAGTCTCAACTCTGTCCACCGAGCTGGAGTGCAGT 660

601 TTTTATTTTAAATTTTGGAGCAGAGTCTCAACTCTGTCCACCGAGCTGGAGTGCAGT 660

661 GCGCAATCTTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTATTTCTCTCTCCCA 720

661 GCGCAATCTTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTATTTCTCTCTCCCA 720

721 GCCTCTGAGTAGTGGGACTACAGGCGCCCAACGCTAGCTAATTTTGTATTTT 780

721 GCCTCTGAGTAGTGGGACTACAGGCGCCCAACGCTAGCTAATTTTGTATTTT 780

781 TAGTAGAGATGGGGTTCAACATGTTTCGCGAGTGTGATCTCTGAGCTTGTGATC 840

781 TAGTAGAGATGGGGTTCAACATGTTTCGCGAGTGTGATCTCTGAGCTTGTGATC 840

841 TGCTGCTCTCGGCTCCCAAGTGTGGGATTTACAGGCGGTGAGCCACCAACGCTGGCTTA 900

841 TGCTGCTCTCGGCTCCCAAGTGTGGGATTTACAGGCGGTGAGCCACCAACGCTGGCTTA 900

901 TTTTAAATTTTGTGTTGAAATGGAATCTCACTCTGTATACCGAGCTGGAGTGCAT 960

901 TTTTAAATTTTGTGTTGAAATGGAATCTCACTCTGTATACCGAGCTGGAGTGCAT 960

961 GCGCAATCTCGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGGATTTCTCTCTCTCA 1020

961 GCGCAATCTCGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGGATTTCTCTCTCTCA 1020

1021 GCCTCCCAAGCAGCTGGGATTTACGGGCACTGCGCCACCAACGCTGGCTTAATTTTGTATTT 1080

1021 GCCTCCCAAGCAGCTGGGATTTACGGGCACTGCGCCACCAACGCTGGCTTAATTTTGTATTT 1080

1081 TCAATTAGAGCGGGTTTTCACATATTTGTGAGGCTGTCTCAAACTCTGAGCTCAGGT 1140

1081 TCAATTAGAGCGGGTTTTCACATATTTGTGAGGCTGTCTCAAACTCTGAGCTCAGGT 1140

1141 GACCACTCTGCTCAGCTTCCAAAGTCTGGGATTTACAGGCGTGGAGCTCAGCTCAGCCAG 1200

1141 GACCACTCTGCTCAGCTTCCAAAGTCTGGGATTTACAGGCGTGGAGCTCAGCTCAGCCAG 1200

1201 CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGTCTTGTATTTGCGCCAGGT 1260

1201 CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGTCTTGTATTTGCGCCAGGT 1260

1261 GGTCTCAAACTCTGGCTTCAATGCAATCTTCCAAATGAGGCAACACCGAGCCAGTCA 1320

1261 GGTCTCAAACTCTGGCTTCAATGCAATCTTCCAAATGAGGCAACACCGAGCCAGTCA 1320

1321 CATTTTAAACAGTTACATCTTTTATTTTATTTAGTATAAGAAAGTAAATAAATAAATGT 1380

1321 CATTTTAAACAGTTACATCTTTTATTTTATTTAGTATAAGAAAGTAAATAAATAAATGT 1380

1381 CAACTCTCAAAATTCAGTAGTAACAGAGTCTTTTATTAACCTTTTAAACAAAGCTTTAGAG 1440

1381 CAACTCTCAAAATTCAGTAGTAACAGAGTCTTTTATTAACCTTTTAAACAAAGCTTTAGAG 1440

1441 CA 1442

1441 CA 1442

RESULT 4

US-09-872-968-1  
Sequence 1, Application US/09872968  
Publication No. US20030050262A1  
GENERAL INFORMATION:  
APPLICANT: Wands, Jack R  
APPLICANT: de la Monte, Suzanne M  
TITLE OF INVENTION: Inhibition of Neurodegeneration  
FILE REFERENCE: 21486-047  
CURRENT APPLICATION NUMBER: US/09/872,968  
CURRENT FILING DATE: 2001-06-01  
NUMBER OF SEQ ID NOS: 2

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1

LENGTH: 1442

TYPE: DNA

ORGANISM: Homo sapiens

US-09-872-968-1

Query Match 100.0%; Score 1442; DB 10; Length 1442;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTATTTTGTGAGATGAGATTTTCTGCTCTGTTGTCCTGAGGCTGAGATGCAATGCGCAA 60  
 DB 1 TTTTATTTTGTGAGATGAGATTTTCTGCTCTGTTGTCCTGAGGCTGAGATGCAATGCGCAA 60  
 QY 61 TCTCAGCTCAGCGCAACCTCCGCTCCCGGCTTCAAGCGATTTCTCTGCTCAGCCTCC 120  
 DB 61 TCTCAGCTCAGCGCAACCTCCGCTCCCGGCTTCAAGCGATTTCTCTGCTCAGCCTCC 120  
 QY 121 CAGTAGCTGGGATTCAGGCAATGTGCAACGCTCGGCTAAATTTTGTATTTTGTATTTAG 180  
 DB 121 CAGTAGCTGGGATTCAGGCAATGTGCAACGCTCGGCTAAATTTTGTATTTTGTATTTAG 180  
 QY 181 TAGAGATGAGATTTCTCCATTTGTTGTCAGGCTGCTCGAACTCCCGACCTCAGATGATC 240  
 DB 181 TAGAGATGAGATTTCTCCATTTGTTGTCAGGCTGCTCGAACTCCCGACCTCAGATGATC 240  
 QY 241 CCTCGCTCCGCTCCCGCTCCCGGCTTCAAGCGATTTCTCTGCTCAGCCTCC 300  
 DB 241 CCTCGCTCCGCTCCCGCTCCCGGCTTCAAGCGATTTCTCTGCTCAGCCTCC 300  
 QY 301 TGGCTAAATTTTGTGTTGAGAAACAGGTTTCTGATGTGCTGCTGCTGCTGCTGCTGCTG 360  
 DB 301 TGGCTAAATTTTGTGTTGAGAAACAGGTTTCTGATGTGCTGCTGCTGCTGCTGCTGCTG 360  
 QY 361 TCAGCAGTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCC 420  
 DB 361 TCAGCAGTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCC 420  
 QY 421 CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 480  
 DB 421 CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 480  
 QY 481 TGCAGTGTGTGATCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 540  
 DB 481 TGCAGTGTGTGATCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCT 540  
 QY 541 CTTAGCTCCCAAGTAGCTGGGACCAAGACATGACCACTACCACTACCACTACCACTACCA 600  
 DB 541 CTTAGCTCCCAAGTAGCTGGGACCAAGACATGACCACTACCACTACCACTACCACTACCA 600  
 QY 601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 660  
 DB 601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 660  
 QY 661 GCGCAATCTGGCTCAGTCACTGCACTCTGCTCCCGGTTCAAGTTATTTCTCTGCTGCT 720  
 DB 661 GCGCAATCTGGCTCAGTCACTGCACTCTGCTCCCGGTTCAAGTTATTTCTCTGCTGCT 720  
 QY 721 GCTCTCTGATGATGAGTCAAGGCTCCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 780  
 DB 721 GCTCTCTGATGATGAGTCAAGGCTCCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 780  
 QY 781 TAGTAGAGATGGGTTTCCATGTTGCGCAGGTTGATCTTGTATCTTGTGATCTGATC 840  
 DB 781 TAGTAGAGATGGGTTTCCATGTTGCGCAGGTTGATCTTGTATCTTGTGATCTTGTGATC 840  
 QY 841 TGCCTGCTCCGCTCCCGCTCCCGGCTTCAAGGCTGAGGATTTACAGGCTGAGGCTGAG 900  
 DB 841 TGCCTGCTCCGCTCCCGCTCCCGGCTTCAAGGCTGAGGATTTACAGGCTGAGGCTGAG 900  
 QY 901 TTTTATTTTATTTTGTGTTGAAATGGAATCTCCTCTGTTACCGAGCTGGAGTGCAT 960  
 DB 901 TTTTATTTTATTTTGTGTTGAAATGGAATCTCCTCTGTTACCGAGCTGGAGTGCAT 960

RESULT 5

US-09-964-678A-1  
 ; Sequence 1, Application US/09964678A  
 ; Publication No. US20030066097A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wands, Jack R.  
 ; APPLICANT: de la Monte, Suzanne  
 ; TITLE OF INVENTION: Transgenic Animals and Cell Lines for Screening Drugs  
 ; TITLE OF INVENTION: Effective for the Treatment or Prevention of  
 ; TITLE OF INVENTION: Alzheimer's Disease  
 ; FILE REFERENCE: 0609.4370002  
 ; CURRENT APPLICATION NUMBER: US/09/964,678A  
 ; CURRENT FILING DATE: 2001-09-28  
 ; PRIOR APPLICATION NUMBER: 09/380,203  
 ; PRIOR FILING DATE: 2000-04-25  
 ; PRIOR APPLICATION NUMBER: PCT/US98/03685  
 ; PRIOR FILING DATE: 1998-02-26  
 ; PRIOR APPLICATION NUMBER: 60/038,908  
 ; PRIOR FILING DATE: 1997-02-26  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 1  
 ; LENGTH: 1442  
 ; TYPE: DNA  
 ; ORGANISM: Unknown  
 ; FEATURE:  
 ; OTHER INFORMATION: AD7c-NTP cDNA  
 ; NAME/KEY: CDS  
 ; LOCATION: (15)..(1139)  
 ; OTHER INFORMATION:  
 ; US-09-964-678A-1

Query Match 100.0%; Score 1442; DB 10; Length 1442;  
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	TTTTTTTTTTTGGAGATGGAGTTTTCGCTCTTGTTGCCAGGCTGGAGTGCATATGCGCAA	60
Db	1	TTTTTTTTTTTGGAGATGGAGTTTTCGCTCTTGTTGCCAGGCTGGAGTGCATATGCGCAA	60
QY	61	TCCTCAGCTCACGCCAACCTCCGCTCCGCGTTCAAGCGATTCTCCTGCCTCAGCTCC	120
Db	61	TCCTCAGCTCACGCCAACCTCCGCTCCGCGTTCAAGCGATTCTCCTGCCTCAGCTCC	120
QY	121	CAGTAGCTGGGATTAACAGCATGTGCAACCCAGCTCGGCTAAATTTGTATTTTTTTAG	180
Db	121	CAGTAGCTGGGATTAACAGCATGTGCAACCCAGCTCGGCTAAATTTGTATTTTTTTAG	180
QY	181	TAGAGATGAGATTCTCCATGTTGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC	240
Db	181	TAGAGATGAGATTCTCCATGTTGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC	240
QY	241	CCTCGCTCTCGGCTCCCAAGTGTCTAGATACAGGACTGGCCACCATGCCGCTCTGCC	300
Db	241	CCTCGCTCTCGGCTCCCAAGTGTCTAGATACAGGACTGGCCACCATGCCGCTCTGCC	300
QY	301	TGGCTAAATTTTGTGTGAAACAGGGTTTCACTGATGTGCCAAGCTGCTCCTGAGC	360
Db	301	TGGCTAAATTTTGTGTGAAACAGGGTTTCACTGATGTGCCAAGCTGCTCCTGAGC	360
QY	361	TCAGGAGTGCACCTGCTCAGCTCCCAAGTGTCTAGATACAGGACTGGCCACCATGCCGCTG	420
Db	361	TCAGGAGTGCACCTGCTCAGCTCCCAAGTGTCTAGATACAGGACTGGCCACCATGCCGCTG	420
QY	421	CTGGCTCTTTTATTTTATTTTATTTTAAAGACAGGTGTCCTCACTCTTACCCAGGATGAG	480
Db	421	CTGGCTCTTTTATTTTATTTTATTTTAAAGACAGGTGTCCTCACTCTTACCCAGGATGAG	480
QY	481	TGCAGTGTGTGATCACAGCTCACTGCAGCTTCAACTCCTGAGATCAAGCATCTCTCTG	540
Db	481	TGCAGTGTGTGATCACAGCTCACTGCAGCTTCAACTCCTGAGATCAAGCATCTCTCTG	540
QY	541	CCTCAGCTCCCAAGTAGCTGGGACCAAGACATGACACCACTACACCTGGCTTAATTTTA	600
Db	541	CCTCAGCTCCCAAGTAGCTGGGACCAAGACATGACACCACTACACCTGGCTTAATTTTA	600
QY	601	TTTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT	660
Db	601	TTTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT	660
QY	661	GGCGCAATCTGGCTCACTGCAACCTCTGCTCCGCGTTCAAGTTATTTCTCTGCCCA	720
Db	661	GGCGCAATCTGGCTCACTGCAACCTCTGCTCCGCGTTCAAGTTATTTCTCTGCCCA	720
QY	721	GGCTCTGAGTAGCTGGGACTACAGCGGCCGACCGCTAGCTAAATTTTGTATTTT	780
Db	721	GGCTCTGAGTAGCTGGGACTACAGCGGCCGACCGCTAGCTAAATTTTGTATTTT	780
QY	781	TAGTAGAGATGGGTTTACCATGTTCCGACAGTTGATCTTGATCTCTGACCTTGATC	840
Db	781	TAGTAGAGATGGGTTTACCATGTTCCGACAGTTGATCTTGATCTCTGACCTTGATC	840
QY	841	TGCTGCTCCTCGGCTCCCAAGTGTGGATTAACAGGCTGAGCCACCAACCGCGCTTA	900
Db	841	TGCTGCTCCTCGGCTCCCAAGTGTGGATTAACAGGCTGAGCCACCAACCGCGCTTA	900
QY	901	TTTTTAAATTTTGTGTTTGTGTTGAAATGGAATCTCACTCTGTTACCGGCTGGAGTGC	960
Db	901	TTTTTAAATTTTGTGTTTGTGTTGAAATGGAATCTCACTCTGTTACCGGCTGGAGTGC	960
QY	961	GGCCAAATCTCGGCTCACTGCAACCTCTGCTCCGCGCTCAAGGATTTCTCTGTCTCA	1020
Db	961	GGCCAAATCTCGGCTCACTGCAACCTCTGCTCCGCGCTCAAGGATTTCTCTGTCTCA	1020
QY	1021	GGCTCCCAAGCAGCTGGGATTAACAGGCTGAGCCACCAACCGCGCTTAATTTTGTATTT	1080
Db	1021	GGCTCCCAAGCAGCTGGGATTAACAGGCTGAGCCACCAACCGCGCTTAATTTTGTATTT	1080

QY	1081	TCATTTAGAGCGGGGTTTTCACCATATTTGTTCAGGCTGGTCTCAAACTCTCTGACCTCAGGT	1140
Db	1081	TCATTTAGAGCGGGGTTTTCACCATATTTGTTCAGGCTGGTCTCAAACTCTCTGACCTCAGGT	1140
QY	1141	GACCCACCTGCTCAGCTTCCAAAGTCTCGGATTTACAGGCTGAGCCACCTCACCCAG	1200
Db	1141	GACCCACCTGCTCAGCTTCCAAAGTCTCGGATTTACAGGCTGAGCCACCTCACCCAG	1200
QY	1201	CGGCTAAATTTAGATAAAAAAATATGTAGCAATGGGGGTTCTTGCTATTTGCCAGGCT	1260
Db	1201	CGGCTAAATTTAGATAAAAAAATATGTAGCAATGGGGGTTCTTGCTATTTGCCAGGCT	1260
QY	1261	GGTCTCAAACTTGGCTTCATGCAATCCTTCCAAATGAGCCACCAACCCAGCCAGTCA	1320
Db	1261	GGTCTCAAACTTGGCTTCATGCAATCCTTCCAAATGAGCCACCAACCCAGCCAGTCA	1320
QY	1321	CATTTTTTAAACAGTTTACATCTTTATTTTAGTATATAAGAAAGTAAATACAATAAATG	1380
Db	1321	CATTTTTTAAACAGTTTACATCTTTATTTTAGTATATAAGAAAGTAAATACAATAAATG	1380
QY	1381	CAAACTGCAAAATTCAGTAGTAACAGAGTTCTTTTATAACTTTTAAACAAAGCTTTAGAG	1440
Db	1381	CAAACTGCAAAATTCAGTAGTAACAGAGTTCTTTTATAACTTTTAAACAAAGCTTTAGAG	1440
QY	1441	CA 1442	
Db	1441	CA 1442	

RESULT 6

US-10-146-130-1  
; Sequence 1, Application US/10146130  
; Publication No. US20030004107A1  
; GENERAL INFORMATION:  
; APPLICANT: AVERBACK, PAUL  
; TITLE OF INVENTION: METHOD OF PREVENTING CELL DEATH USING SEGMENTS OF  
; FILE REFERENCE: 59003.000007  
; CURRENT APPLICATION NUMBER: US/10/146,130  
; CURRENT FILING DATE: 2002-08-06  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (15)..(1139)  
US-10-146-130-1

Query Match 100.0%; Score 1442; DB 14; Length 1442;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TTTTTTTTTTTGGAGATGGAGTTTTCGCTCTTGTTGCCAGGCTGGAGTGCATATGCGCAA	60
Db	1	TTTTTTTTTTTGGAGATGGAGTTTTCGCTCTTGTTGCCAGGCTGGAGTGCATATGCGCAA	60
QY	61	TCTCAGCTCACGCCAACCTCCGCTCCGCGTTCAAGCGATTCTCCTGCCTCAGCTCC	120
Db	61	TCTCAGCTCACGCCAACCTCCGCTCCGCGTTCAAGCGATTCTCCTGCCTCAGCTCC	120
QY	121	CAGTAGCTGGGATTAACAGCATGTGCAACCCAGCTCGGCTAAATTTGTATTTTTTTAG	180
Db	121	CAGTAGCTGGGATTAACAGCATGTGCAACCCAGCTCGGCTAAATTTGTATTTTTTTAG	180
QY	181	TAGAGATGAGATTCTCCATGTTGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC	240
Db	181	TAGAGATGAGATTCTCCATGTTGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC	240
QY	241	CCTCGCTCTCGGCTCCCAAGTGTAGATACAGGACTGGCCACCAATGCCCGGCTCTGCC	300

Db 241 CCTCGCTCGGCTCCCAAAGTGTAGATACAGGACTGGCCACCATGCGGCTCTGCC 300  
Qy 301 TGGCTAAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCCCAGCTGGTCTCTGAGC 360  
Db 301 TGGCTAAATTTTGTGTAGAAACAGGGTTTCACTGATGTGCCCCAGCTGGTCTCTGAGC 360  
Qy 361 TCAAGCAGTCCACCTGCTCCCTCAGCCTCCCAAAGTGTGCGATTTACAGGCGTGCAGCGTGC 420  
Db 361 TCAAGCAGTCCACCTGCTCCCTCAGCCTCCCAAAGTGTGCGATTTACAGGCGTGCAGCGTGC 420  
Qy 421 CTGGCCCTTTTATTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG 480  
Db 421 CTGGCCCTTTTATTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG 480  
Qy 481 TGCAGTGGTGTATCACGCTCACTGACGCTTCACTCTGAGATCAAGCATCTCTCTG 540  
Db 481 TGCAGTGGTGTATCACGCTCACTGACGCTTCACTCTGAGATCAAGCATCTCTCTG 540  
Qy 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGCTAAATTTTA 600  
Db 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGCTAAATTTTA 600  
Qy 601 TTTTATTTTAAATTTTGTAGACAGAGTCTCACTCTCTCACCAGGCTGAGTGCAGT 660  
Db 601 TTTTATTTTAAATTTTGTAGACAGAGTCTCACTCTCTCACCAGGCTGAGTGCAGT 660  
Qy 661 GGGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTATTTCTCTGCCCA 720  
Db 661 GGGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTATTTCTCTGCCCA 720  
Qy 721 GCTCTCTGAGTAGCTGGGACTACAGGCGCCACACGCTAGCTAATTTTGTATTTT 780  
Db 721 GCTCTCTGAGTAGCTGGGACTACAGGCGCCACACGCTAGCTAATTTTGTATTTT 780  
Qy 781 TAGTAGAGTGGGTTCACTATTTGCCAGGTGATCTTGATCTCTGACCTTGATGATC 840  
Db 781 TAGTAGAGTGGGTTCACTATTTGCCAGGTGATCTTGATCTCTGACCTTGATGATC 840  
Qy 841 TGCCTGCTCGGCTCCCAAAGTGTGGATTTACAGGCGTGAGCCACCGCCGCTTA 900  
Db 841 TGCCTGCTCGGCTCCCAAAGTGTGGATTTACAGGCGTGAGCCACCGCCGCTTA 900  
Qy 901 TTTTAAATTTTGTGTTGAAATGGAATCTCACTCTGTATACCCAGGCTGAGTGCAT 960  
Db 901 TTTTAAATTTTGTGTTGAAATGGAATCTCACTCTGTATACCCAGGCTGAGTGCAT 960  
Qy 961 GGGCAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTTCTCTGTCTCA 1020  
Db 961 GGGCAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTTCTCTGTCTCA 1020  
Qy 1021 GCTTCCCAAGCAGCTGGGATTAAGGCACTGCGCACCTGCGGCTCAAGCGATTTTGTATTT 1080  
Db 1021 GCTTCCCAAGCAGCTGGGATTAAGGCACTGCGCACCTGCGGCTCAAGCGATTTTGTATTT 1080  
Qy 1081 TCATTAGAGCGGGTTTCAACATATTTGTCAAGCTGTGCTCAAACTCTCTGACCTCAGT 1140  
Db 1081 TCATTAGAGCGGGTTTCAACATATTTGTCAAGCTGTGCTCAAACTCTCTGACCTCAGT 1140  
Qy 1141 GACCCACCTGCTCAGCCTTCCAAAGTGTGCGATTTACAGGCGTGAGCCACCTCAAG 1200  
Db 1141 GACCCACCTGCTCAGCCTTCCAAAGTGTGCGATTTACAGGCGTGAGCCACCTCAAG 1200  
Qy 1201 CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGTCTTGTATGTGCGCAGGCT 1260  
Db 1201 CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGTCTTGTATGTGCGCAGGCT 1260  
Qy 1261 GGTCTCAAACTTCTGGCTTCAATGCAATCTTCAAAATGAGCCACCAACCCAGCAGTCA 1320  
Db 1261 GGTCTCAAACTTCTGGCTTCAATGCAATCTTCAAAATGAGCCACCAACCCAGCAGTCA 1320  
Qy 1321 CATTTTAAACAGTTTACATCTTTTATTTTAGTATACCTAGAAAGTAAATACAATAAATGT 1380

Db 1321 CATTTTAAACAGTTTACATCTTTATTTTAGTATACCTAGAAAGTAAATACAATAAATGT 1380  
Qy 1381 CAAACCTGCAATTTCACTAGTAGTAACAGAGTCTTTTATTAACCTTTTAAACAAAGCTTTAGAG 1440  
Db 1381 CAAACCTGCAATTTCACTAGTAGTAACAGAGTCTTTTATTAACCTTTTAAACAAAGCTTTAGAG 1440  
Qy 1441 CA 1442  
Db 1441 CA 1442

RESULT 7

US-10-092-934-1  
; Sequence 1, Application US/10092934  
; Publication No. US20030054990A1  
; GENERAL INFORMATION:  
; APPLICANT: AVERBACK, PAUL  
; TITLE OF INVENTION: METHODS OF USING NEURAL THREAD PROTEINS TO TREAT TUMORS  
; TITLE OF INVENTION: AND CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF  
; FILE REFERENCE: 018792-0199  
; CURRENT APPLICATION NUMBER: US/10/092,934  
; CURRENT FILING DATE: 2002-06-05  
; PRIOR APPLICATION NUMBER: 60/273,957  
; PRIOR FILING DATE: 2001-03-08  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (15)..(1139)  
US-10-092-934-1

Query Match 100.0%; Score 1442; DB 14; Length 1442;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TTTTATTTTGTGAGATGGAGTTTTCGCTCTTGTGTTGCCAGGCTGAGTGCATATGGCGCAA 60  
Db 1 TTTTATTTTGTGAGATGGAGTTTTCGCTCTTGTGTTGCCAGGCTGAGTGCATATGGCGCAA 60  
Qy 61 TCTCAGCTCACGCAACCTCCGCTCCCGGGTTCAAGGATTTCTCTGCTCAGCTCCC 120  
Db 61 TCTCAGCTCACGCAACCTCCGCTCCCGGGTTCAAGGATTTCTCTGCTCAGCTCCC 120  
Qy 121 CAGTAGCTGGGATTTACAGGCAATGTGCACCCAGCTCGGCTAATTTTGTATTTTGTAG 180  
Db 121 CAGTAGCTGGGATTTACAGGCAATGTGCACCCAGCTCGGCTAATTTTGTATTTTGTAG 180  
Qy 181 TAGAGATGGAGTTTCTCATGTTGTGTCAGGCTGCTGCTGAACTCCGACCTCAGATGATC 240  
Db 181 TAGAGATGGAGTTTCTCATGTTGTGTCAGGCTGCTGCTGAACTCCGACCTCAGATGATC 240  
Qy 241 CCTCGCTCTCGGCTCCCAAAGTGTAGTACAGGACTGGCCACCATGCCCGGCTCTGCC 300  
Db 241 CCTCGCTCTCGGCTCCCAAAGTGTAGTACAGGACTGGCCACCATGCCCGGCTCTGCC 300  
Qy 301 TGGCTAAATTTTGTGTTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGC 360  
Db 301 TGGCTAAATTTTGTGTTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGC 360  
Qy 361 TCAAGCAGTCCACCTGCTCAGCCTCCCAAAGTGTGCGATTTACAGGCGTGCAGCGTGC 420  
Db 361 TCAAGCAGTCCACCTGCTCAGCCTCCCAAAGTGTGCGATTTACAGGCGTGCAGCGTGC 420  
Qy 421 CTGGCTTTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG 480  
Db 421 CTGGCTTTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG 480  
Qy 481 TGCAGTGGTGTATCACAGCTCACCTGACGCTTCACTCTGAGATCAAGCATCTCTCTG 540

```
Db 481 TGCAAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCTGAGATCAAGCATCTCTCTG 540
Qy 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGACCACTACACCTGGCTGAATTTTAA 600
Db 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGACCACTACACCTGGCTGAATTTTAA 600
Qy 601 TTTTATTTTAAATTTTGGAGACAGAGTCTCAACTCTCTGACCAAGCTGAGTGCAAT 660
Db 601 TTTTATTTTAAATTTTGGAGACAGAGTCTCAACTCTCTGACCAAGCTGAGTGCAAT 660
Qy 661 GCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTTATTTCTCTGCCCA 720
Db 661 GCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTTATTTCTCTGCCCA 720
Qy 721 GCTCTCTAGTAGTGGGACTACAGGCGCCACACGCTAGCTAATTTTGTATTTT 780
Db 721 GCTCTCTAGTAGTGGGACTACAGGCGCCACACGCTAGCTAATTTTGTATTTT 780
Qy 781 TAGTAGAGATGGGTTTCAACATGTCGCCAGGTTGATCTTGATCTCTGGACCTTGTATC 840
Db 781 TAGTAGAGATGGGTTTCAACATGTCGCCAGGTTGATCTTGATCTCTGGACCTTGTATC 840
Qy 841 TGCCTGCTTGGCTCCCAAGTGTCTGGATTTACAGGCGTGAGCCACACGCGCCGCTTA 900
Db 841 TGCCTGCTTGGCTCCCAAGTGTCTGGATTTACAGGCGTGAGCCACACGCGCCGCTTA 900
Qy 901 TTTTAAATTTTGTGTTGAAATGGAATCTCACTCTGTTTACCAGGCTGGAGTGCAAT 960
Db 901 TTTTAAATTTTGTGTTGAAATGGAATCTCACTCTGTTTACCAGGCTGGAGTGCAAT 960
Qy 961 GCGCAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATCTCTCTCTCA 1020
Db 961 GCGCAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATCTCTCTCTCA 1020
Qy 1021 GCTCTCCCAAGAGCTGGGATTTACGGGCACTTCCGACCAACACCGCTTAATTTTGTATTT 1080
Db 1021 GCTCTCCCAAGAGCTGGGATTTACGGGCACTTCCGACCAACACCGCTTAATTTTGTATTT 1080
Qy 1081 TCATTAGAGCGGGGTTTCAACATATTTGTCTAGGCTGGTCTCAAACTCTGACCTCAGGT 1140
Db 1081 TCATTAGAGCGGGGTTTCAACATATTTGTCTAGGCTGGTCTCAAACTCTGACCTCAGGT 1140
Qy 1141 GACCCACCTGCTCAGCCTTCCAAAGTCTGGGATTTACAGGCGTGAGCCACCTCACCCAG 1200
Db 1141 GACCCACCTGCTCAGCCTTCCAAAGTCTGGGATTTACAGGCGTGAGCCACCTCACCCAG 1200
Qy 1201 CCGGCTAATTTAGATAAAAAAATATGTAGCAATGGGGGCTCTTGCTATGTTGCCAGGCT 1260
Db 1201 CCGGCTAATTTAGATAAAAAAATATGTAGCAATGGGGGCTCTTGCTATGTTGCCAGGCT 1260
Qy 1261 GGTCTCAAACTTCTGGCTTCAATGCAATCTTCCAAATGAGCCACCAACCCAGCAGTCA 1320
Db 1261 GGTCTCAAACTTCTGGCTTCAATGCAATCTTCCAAATGAGCCACCAACCCAGCAGTCA 1320
Qy 1321 CATTTTTHAAACAGTTACATCTTTTATTTAGTATAGTAAAGTAAATACATAAATATGT 1380
Db 1321 CATTTTTHAAACAGTTACATCTTTTATTTAGTATAGTAAAGTAAATACATAAATATGT 1380
Qy 1381 CAAACTGCAAAATTCAGTAGTAAACAGAGTCTTTTATTAATCTTTTAAACAAAGCTTTAG 1440
Db 1381 CAAACTGCAAAATTCAGTAGTAAACAGAGTCTTTTATTAATCTTTTAAACAAAGCTTTAG 1440
Qy 1441 CA 1442
Db 1441 CA 1442
```

RESULT 8  
US-10-153-334-53  
; Sequence 53, Application US/10153334  
; Publication No. US20030096350A1  
; GENERAL INFORMATION:

; APPLICANT: AVERBACK, PAUL  
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER  
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF  
; TITLE OF INVENTION: CELLS  
; FILE REFERENCE: 59003-000006  
; CURRENT APPLICATION NUMBER: US/10/153,334  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: 60/293,156  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 53  
; LENGTH: 1442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (15)...(1139)  
US-10-153-334-53

Query Match 100.0%; Score 1442; DB 14; Length 1442;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 TTTTATTTTGTGAGATGGAGTTTTCGCTCTTGTGTCGCCAGGCTGGAGTGCAATGGCGCA 60  
Db 1 TTTTATTTTGTGAGATGGAGTTTTCGCTCTTGTGTCGCCAGGCTGGAGTGCAATGGCGCA 60  
Qy 61 TCTCAGCTCACCGCAACCTCCGCGCTCCCGGGTTCGAAGCGATTTCTCTGCCCTCAGCCTCCC 120  
Db 61 TCTCAGCTCACCGCAACCTCCGCGCTCCCGGGTTCGAAGCGATTTCTCTGCCCTCAGCCTCCC 120  
Qy 121 CAGTAGCTGGGATTTACAGGCGATGTGCAACCGCTCGGCTAATTTGTATTTTGTATTTT 180  
Db 121 CAGTAGCTGGGATTTACAGGCGATGTGCAACCGCTCGGCTAATTTGTATTTTGTATTTT 180  
Qy 181 TAGAGATGGAGTTTCTCCATGTTGGTCTAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240  
Db 181 TAGAGATGGAGTTTCTCCATGTTGGTCTAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240  
Qy 241 CCTCGCTCTCGGCTTCCCAAGTGTAGATACAGGACTGGGCCACCATGCCCGGCTCTGCCC 300  
Db 241 CCTCGCTCTCGGCTTCCCAAGTGTAGATACAGGACTGGGCCACCATGCCCGGCTCTGCCC 300  
Qy 301 TGGCTAATTTTGTGTTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTCTGAGC 360  
Db 301 TGGCTAATTTTGTGTTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTCTGAGC 360  
Qy 361 TCAGCAGTCCACCTGCTCAGCCTCCCAAGTGTCTGGGATTTACAGGCGTGAGCGCGTGC 420  
Db 361 TCAGCAGTCCACCTGCTCAGCCTCCCAAGTGTCTGGGATTTACAGGCGTGAGCGCGTGC 420  
Qy 421 CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480  
Db 421 CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480  
Qy 481 TGCAGTGGTGTGATCAGAGCTCACTGAGCCTTCAACTCTCTGAGATCAAGCATCTCTCTG 540  
Db 481 TGCAGTGGTGTGATCAGAGCTCACTGAGCCTTCAACTCTCTGAGATCAAGCATCTCTCTG 540  
Qy 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGACCACTACACCTGGGCTAATTTTAA 600  
Db 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGACCACTACACCTGGGCTAATTTTAA 600  
Qy 601 TTTTATTTTAAATTTTGTGAGACAGAGTCTCAACTCTGTCGCCAGGCTGAGTGCAAT 660  
Db 601 TTTTATTTTAAATTTTGTGAGACAGAGTCTCAACTCTGTCGCCAGGCTGAGTGCAAT 660  
Qy 661 GCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTTATTTCTCTGCCCA 720  
Db 661 GCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTTATTTCTCTGCCCA 720  
Qy 721 GCCTCTCTAGTAGTGGGACTACAGGCGCCACACGCTAGCTAATTTTGTATTTT 780  
Db 721 GCCTCTCTAGTAGTGGGACTACAGGCGCCACACGCTAGCTAATTTTGTATTTT 780

```

721 GCCTCTGAGTAGCTGGGACTACAGGCGCCACCGCCTAGCTAAATTTTGTATTTT 780
781 TAGTAGAGATGGGGTTTACCAGTCTCGCCAGGTTTGTATCTTGATCTCTGACCTTGTGATC 840
781 TAGTAGAGATGGGGTTTACCAGTCTCGCCAGGTTTGTATCTTGATCTCTGACCTTGTGATC 840
841 TGCCCTGCTCGGCTCCCCAAAGTGTGGGATTCAGGCGTGAGCCACCGCCGCTTCA 900
841 TGCCCTGCTCGGCTCCCCAAAGTGTGGGATTCAGGCGTGAGCCACCGCCGCTTCA 900
901 TTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAT 960
901 TTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAT 960
961 GGCCTGCTCGGCTCACTGCAACCTCTGCTCCCGGCTCAAGCGATTTCTCTGCTCA 1020
961 GGCCTGCTCGGCTCACTGCAACCTCTGCTCCCGGCTCAAGCGATTTCTCTGCTCA 1020
1021 GGCCTGCTCGGCTCACTGCAACCTCTGCTCCCGGCTCAAGCGATTTCTCTGCTCA 1080
1021 GGCCTGCTCGGCTCACTGCAACCTCTGCTCCCGGCTCAAGCGATTTCTCTGCTCA 1080
1081 TCATTAGAGCGGGTTTACCAATATTTGTCAGGCTGGTCTCAAACTCTCTGACCTCAGGT 1140
1081 TCATTAGAGCGGGTTTACCAATATTTGTCAGGCTGGTCTCAAACTCTCTGACCTCAGGT 1140
1141 GACCCACCTGCTGACCTTCCAAAGTGTGGGATTTACAGGCGTGAGCCACCTCA 1200
1141 GACCCACCTGCTGACCTTCCAAAGTGTGGGATTTACAGGCGTGAGCCACCTCA 1200
1201 CGGGCTAAATTTAGATAAAATATATAGCAATGGGGGCTCTTGCTATGTGCGCCAGGCT 1260
1201 CGGGCTAAATTTAGATAAAATATATAGCAATGGGGGCTCTTGCTATGTGCGCCAGGCT 1260
1261 GGTCTCAAACTTCTGGCTTCAAGCAATCTTCCAAATGAGCCACCAACACCGCCAGTCA 1320
1261 GGTCTCAAACTTCTGGCTTCAAGCAATCTTCCAAATGAGCCACCAACACCGCCAGTCA 1320
1321 CAATTTTAAACAGTTACATCTTTATTTAGTATCTAGAAAGTAAATCAATAAATCATGT 1380
1321 CAATTTTAAACAGTTACATCTTTATTTAGTATCTAGAAAGTAAATCAATAAATCATGT 1380
1381 CAAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATAACTTTTAAACAAAGCTTTAGAG 1440
1381 CAAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATAACTTTTAAACAAAGCTTTAGAG 1440
1441 CA 1442
1441 CA 1442

```

RESULT 9  
 US-10-198-069-48  
 ; Sequence 48, Application US/10198069  
 ; Publication No. US20030096756A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: AVERBACK, PAUL  
 ; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER  
 ; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF  
 ; TITLE OF INVENTION: CELLS  
 ; FILE REFERENCE: 59003.000009  
 ; CURRENT APPLICATION NUMBER: US/10/198,069  
 ; CURRENT FILING DATE: 2002-07-19  
 ; PRIOR APPLICATION NUMBER: 60/306,161  
 ; PRIOR FILING DATE: 2001-07-19  
 ; PRIOR APPLICATION NUMBER: 60/306,150  
 ; PRIOR FILING DATE: 2001-07-19  
 ; PRIOR APPLICATION NUMBER: 60/331,477  
 ; PRIOR FILING DATE: 2001-11-16  
 ; NUMBER OF SEQ ID NOS: 48  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 48

```

; LENGTH: 1442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15)..(1139)
US-10-198-069-48

Query Match
Best Local Similarity 100.0%; Score 1442; DB 14; Length 1442;
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTATTTTTCAGATGGAGTTTTCGCTCTTGTGTCAGGCTGGAGTCAATGGCGCAA 60
DB 1 TTTTATTTTTCAGATGGAGTTTTCGCTCTTGTGTCAGGCTGGAGTCAATGGCGCAA 60
QY 61 TCTCAGCTCACGCAACCTCCGCTCCCGGTTCAAGCGATTTCTCTGCTCAGCTCCC 120
DB 61 TCTCAGCTCACGCAACCTCCGCTCCCGGTTCAAGCGATTTCTCTGCTCAGCTCCC 120
QY 121 CAGTAGCTGGGATTCAGGCATGTGCACCGCTCGGCTAAATTTTGTATTTTGTAG 180
DB 121 CAGTAGCTGGGATTCAGGCATGTGCACCGCTCGGCTAAATTTTGTATTTTGTAG 180
QY 181 TAGAGATGGAGTTTCTCCATGTTGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC 240
DB 181 TAGAGATGGAGTTTCTCCATGTTGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATC 240
QY 241 CTTCCGCTCTCGGCTCCCAAGTGTAGATACAGGACTGGCCACCAATGCCCCGCTCTGCC 300
DB 241 CTTCCGCTCTCGGCTCCCAAGTGTAGATACAGGACTGGCCACCAATGCCCCGCTCTGCC 300
QY 301 TGGCTAAATTTTGTGTAGAAACAGGTTTCACTGATGTGCCCAAGCTGCTCTCTGAGC 360
DB 301 TGGCTAAATTTTGTGTAGAAACAGGTTTCACTGATGTGCCCAAGCTGCTCTCTGAGC 360
QY 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGGGATTTACAGGCGTGAGCGCTGC 420
DB 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAGTGTGGGATTTACAGGCGTGAGCGCTGC 420
QY 421 CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480
DB 421 CTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 480
QY 481 TGCAGTGTGTGATCAGAGCTCAGTGTGAGCTTCACTGATCTCTGAGATCAAGCATCTCTCTG 540
DB 481 TGCAGTGTGTGATCAGAGCTCAGTGTGAGCTTCACTGATCTCTGAGATCAAGCATCTCTCTG 540
QY 541 CTTCAGCTCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGCTAAATTTT 600
DB 541 CTTCAGCTCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGCTAAATTTT 600
QY 601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660
DB 601 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 660
QY 661 GCGCAATCTTGGCTCAGTGTGAGCTTCCGCTCCCGGTTCAAGTTATTTCTCTGCCCCA 720
DB 661 GCGCAATCTTGGCTCAGTGTGAGCTTCCGCTCCCGGTTCAAGTTATTTCTCTGCCCCA 720
QY 721 GCCTCTGAGTAGCTGGGACTACAGGCGCCCAACCGCTAGCTAAATTTTGTATTTT 780
DB 721 GCCTCTGAGTAGCTGGGACTACAGGCGCCCAACCGCTAGCTAAATTTTGTATTTT 780
QY 781 TAGTAGAGATGGGGTTTACCAATGTTTGGCAGGTTGATCTTGATCTCTGAGCTTGTGATC 840
DB 781 TAGTAGAGATGGGGTTTACCAATGTTTGGCAGGTTGATCTTGATCTCTGAGCTTGTGATC 840
QY 841 TGCCTGCTCGGCTCCCAAGTGTGGGATTTACAGGCGTGAGCCACCGCCGCTTCA 900
DB 841 TGCCTGCTCGGCTCCCAAGTGTGGGATTTACAGGCGTGAGCCACCGCCGCTTCA 900
QY 901 TTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAT 960

```



Db 901 TTTTAAATTTTGTGTTGTAATGGAATCTCACTCTGTATCCAGGCTGGAGTGCAAT 960  
Qy 961 GGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTCTCTGTCTCA 1020  
Db 961 GGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTCTCTGTCTCA 1020  
Qy 1021 GCTTCCCAAGCAGCTGGGATTAAGGACCTGCGCACCAACCCCGCTAAATTTTGTATTT 1080  
Db 1021 GCTTCCCAAGCAGCTGGGATTAAGGACCTGCGCACCAACCCCGCTAAATTTTGTATTT 1080  
Qy 1081 TCATTAGAGCGCGGGTTTCCACCATATTTGTCAAGCTGGTCTCAAACTCTCTGACCTCAGT 1140  
Db 1081 TCATTAGAGCGCGGGTTTCCACCATATTTGTCAAGCTGGTCTCAAACTCTCTGACCTCAGT 1140  
Qy 1141 GACCCACCTGCTCAGCTTCCAAAGTGTGGGATTAACAGGGGTGAGCCACCTCACCCAG 1200  
Db 1141 GACCCACCTGCTCAGCTTCCAAAGTGTGGGATTAACAGGGGTGAGCCACCTCACCCAG 1200  
Qy 1201 CGGGCTAAATTTAGATAAATAATATGTAGCAATGGGGGTCTTGCTATGTTGCCAGGCT 1260  
Db 1201 CGGGCTAAATTTAGATAAATAATATGTAGCAATGGGGGTCTTGCTATGTTGCCAGGCT 1260  
Qy 1261 GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACACCCAGCAGTCA 1320  
Db 1261 GGTCTCAAACTTCTGGCTTCATGCAATCTTCCAAATGAGCCACACCCAGCAGTCA 1320  
Qy 1321 CAATTTTAAACAGTTACATCTTTATTTTATGATATACTAGAAAGTAATAACAATAAATGT 1380  
Db 1321 CAATTTTAAACAGTTACATCTTTATTTTATGATATACTAGAAAGTAATAACAATAAATGT 1380  
Qy 1381 CAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATTAACCTTTTAAACAGCTTTAG 1440  
Db 1381 CAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATTAACCTTTTAAACAGCTTTAG 1440  
Qy 1441 CA 1442  
Db 1441 CA 1442

RESULT 10  
US-10-198-070-125  
; Sequence 125, Application US/10198070  
; Publication No. US20030109437A1  
; GENERAL INFORMATION:  
; APPLICANT: AVERBACK, PAUL  
; APPLICANT: GEMMELL, JACK  
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER  
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF  
; TITLE OF INVENTION: CELLS  
; FILE REFERENCE: 59003.000008  
; CURRENT APPLICATION NUMBER: US/10/198,070  
; CURRENT FILING DATE: 2002-07-19  
; PRIOR APPLICATION NUMBER: 60/306,161  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: 60/306,150  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: 60/331,477  
; PRIOR FILING DATE: 2001-11-16  
; NUMBER OF SEQ ID NOS: 125  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 125  
; LENGTH: 1442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (15)..(1139)  
US-10-198-070-125

Query Match 100.0%; Score 1442; DB 15; Length 1442;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTTATTTTTCAGATGAGTATTTTCGCTCTCTGTTGCTCCAGGCTGAGTGCATATGGCGAA 60  
Db 1 TTTTATTTTTCAGATGAGTATTTTCGCTCTCTGTTGCTCCAGGCTGAGTGCATATGGCGAA 60  
Qy 61 TCTCAGCTCACCGCAACCTCCGCTCCCGGGTTCAAGCGATTCTCCTGCTCAGCTCC 120  
Db 61 TCTCAGCTCACCGCAACCTCCGCTCCCGGGTTCAAGCGATTCTCCTGCTCAGCTCC 120  
Qy 121 CAGTAGCTGGGATTAACAGGCATGTGCACCCAGCTCGGCTAATTTTGTATTTTGTAG 180  
Db 121 CAGTAGCTGGGATTAACAGGCATGTGCACCCAGCTCGGCTAATTTTGTATTTTGTAG 180  
Qy 181 TAGAGATGGAGTTTCTCAGTTGTGTGCTCAGGCTGGTCTGAACTCCGACCTCAGATGATC 240  
Db 181 TAGAGATGGAGTTTCTCAGTTGTGTGCTCAGGCTGGTCTGAACTCCGACCTCAGATGATC 240  
Qy 241 CCTCCGCTCTCGGCTCCCAAAAGTGTAGATACAGGACTGGCCACCATGCCCCGGCTCTGCC 300  
Db 241 CCTCCGCTCTCGGCTCCCAAAAGTGTAGATACAGGACTGGCCACCATGCCCCGGCTCTGCC 300  
Qy 301 TGGCTAAATTTTGTGTAGAAAACAGGGTTTCACTGATGTGCCAAAGCTGGTCTCCTGAGC 360  
Db 301 TGGCTAAATTTTGTGTAGAAAACAGGGTTTCACTGATGTGCCAAAGCTGGTCTCCTGAGC 360  
Qy 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAAAGTGTGGGATTAACAGGCGTGCAGCGTGC 420  
Db 361 TCAAGCAGTCCACCTGCTCAGCTCCCAAAAGTGTGGGATTAACAGGCGTGCAGCGTGC 420  
Qy 421 CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCCTTACCCAGGATGAAG 480  
Db 421 CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTGTCCTTACCCAGGATGAAG 480  
Qy 481 TGCAGTGTGTGATCAGCAGCTCACTGAGCCCTTCAACTCTTGAGATCAAGCATCTCCTG 540  
Db 481 TGCAGTGTGTGATCAGCAGCTCACTGAGCCCTTCAACTCTTGAGATCAAGCATCTCCTG 540  
Qy 541 CCTCAGCTCCCAAGTAGCTGGACCAAAAGACATGCAACACCTGGCTTAATTTTAA 600  
Db 541 CCTCAGCTCCCAAGTAGCTGGACCAAAAGACATGCAACACCTGGCTTAATTTTAA 600  
Qy 601 TTTTATTTTAAATTTTGTAGACAGAGTCTCAACTGTGTCAACCCAGGCTGGAGTGCAAT 660  
Db 601 TTTTATTTTAAATTTTGTAGACAGAGTCTCAACTGTGTCAACCCAGGCTGGAGTGCAAT 660  
Qy 661 GCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCCTGCCCA 720  
Db 661 GCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCCTGCCCA 720  
Qy 721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCAACAGCTAGCTAATTTTGTATTTT 780  
Db 721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCAACAGCTAGCTAATTTTGTATTTT 780  
Qy 781 TAGTAGAGATGGGGTTTCAACATGTTTCGCCAGTTGATCTTGATCTCTGACCTTGTGATC 840  
Db 781 TAGTAGAGATGGGGTTTCAACATGTTTCGCCAGTTGATCTTGATCTCTGACCTTGTGATC 840  
Qy 841 TGCCTGCTCGGCTCCCAAAAGTGTGGGATTAACAGGCGTGAAGCCACCGCCGGCTTA 900  
Db 841 TGCCTGCTCGGCTCCCAAAAGTGTGGGATTAACAGGCGTGAAGCCACCGCCGGCTTA 900  
Qy 901 TTTTAAATTTTGTGTTGTAATGGAATCTCACTGTGTTACCCAGGCTGGAGTGCAAT 960  
Db 901 TTTTAAATTTTGTGTTGTAATGGAATCTCACTGTGTTACCCAGGCTGGAGTGCAAT 960  
Qy 961 GSCCAAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGGATTTCTCCTGTCTCA 1020  
Db 961 GSCCAAAATCTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGGATTTCTCCTGTCTCA 1020  
Qy 1021 GCCTCCCAAGCAGCTGGGATTAACGGGCACTGCGCAACCCAGCTCAATTTTGTATTT 1080  
Db 1021 GCCTCCCAAGCAGCTGGGATTAACGGGCACTGCGCAACCCAGCTCAATTTTGTATTT 1080



QY 1081 TCATTAGAGCGGGGTTTCAACATATTTGTCAGGCTGGTCTCAAACTCCTGACCTCAGGT 1140  
 Db 1081 TCATTAGAGCGGGGTTTCAACATATTTGTCAGGCTGGTCTCAAACTCCTGACCTCAGGT 1140  
 QY 1141 GACCCACCTGCTCAGCCTTCCAAAGTGTGGGATTTACAGGCGTGAGCCACTCACCAG 1200  
 Db 1141 GACCCACCTGCTCAGCCTTCCAAAGTGTGGGATTTACAGGCGTGAGCCACTCACCAG 1200  
 QY 1201 CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGCTTTGCTATGTGCGCCAGGCT 1260  
 Db 1201 CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGCTTTGCTATGTGCGCCAGGCT 1260  
 QY 1261 GGTCTCAAACTTCTGCTTCAATGCAATCTTCAAAATGAGCCAAACACCCAGTCAC 1320  
 Db 1261 GGTCTCAAACTTCTGCTTCAATGCAATCTTCAAAATGAGCCAAACACCCAGTCAC 1320  
 QY 1321 CATTTTAAACAGTTACATCTTTATTTAGTATATACTAGAAAGTAATACAAATATGT 1380  
 Db 1321 CATTTTAAACAGTTACATCTTTATTTAGTATATACTAGAAAGTAATACAAATATGT 1380  
 QY 1381 CAAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATACTTTTAAACAAAGCTTAGAG 1440  
 Db 1381 CAAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATACTTTTAAACAAAGCTTAGAG 1440  
 QY 1441 CA 1442  
 Db 1441 CA 1442

RESULT 11

US-10-755-889-409  
 ; Sequence 409, Application US/10755889  
 ; Publication No. US20040171823A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bristol-Myers Squibb Company  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB  
 ; FILE OF INVENTION: PATHWAY  
 ; FILE REFERENCE: D0284 NP  
 ; CURRENT APPLICATION NUMBER: US/10/755,889  
 ; CURRENT FILING DATE: 2004-01-13  
 ; PRIOR APPLICATION NUMBER: U.S. 60/440,068  
 ; PRIOR FILING DATE: 2003-01-14  
 ; PRIOR APPLICATION NUMBER: U.S. 60/469,757  
 ; PRIOR FILING DATE: 2003-05-12  
 ; NUMBER OF SEQ ID NOS: 823  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 409  
 ; LENGTH: 1442  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-755-889-409

Query Match 100.0%; Score 1442; DB 19; Length 1442;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TTTTATTTTTCAGATGGAGTTTTCGCTCTGTGTCAGGCTGGAGTGCATGGCGCAA 60  
 Db 1 TTTTATTTTTCAGATGGAGTTTTCGCTCTGTGTCAGGCTGGAGTGCATGGCGCAA 60  
 QY 61 TCTCAGCTCAGCGCAACTCCGCTCCCGGGTTCAAGCGATTTCTCTGCTCAGGCTCC 120  
 Db 61 TCTCAGCTCAGCGCAACTCCGCTCCCGGGTTCAAGCGATTTCTCTGCTCAGGCTCC 120  
 QY 121 CAGTAGCTGGATTTACAGGATGTGCAACCCAGCTCGGCTAAATTTGTAATTTTTTAG 180  
 Db 121 CAGTAGCTGGATTTACAGGATGTGCAACCCAGCTCGGCTAAATTTGTAATTTTTTAG 180  
 QY 181 TAGAGATGAGTTTCTCCATGTTGTCAGGCTGGTCTGCACTCCCGACCTCAGATGATC 240  
 Db 181 TAGAGATGAGTTTCTCCATGTTGTCAGGCTGGTCTGCACTCCCGACCTCAGATGATC 240  
 QY 241 CCTCGCTCTCGGCTCTCCCAAGTGTGTAGATACAGGACTGGCCACCATGCGGCTCTGCC 300

Db 241 CCTCGCTCTCGGCTCTCCCAAGTGTGTAGATACAGGACTGGCCACCATGCGCGCTCTGCC 300  
 QY 301 TGGCTAAATTTTGTGTAGAAACAGGTTTCACTGATGTGCCCAAGCTGTTCTCTGAGC 360  
 Db 301 TGGCTAAATTTTGTGTAGAAACAGGTTTCACTGATGTGCCCAAGCTGTTCTCTGAGC 360  
 QY 361 TCAACAGCTGCCACTGCGCTCAGGCTCCCAAAAGTGTGGGATTTACAGGCGTGAGCGCTGC 420  
 Db 361 TCAACAGCTGCCACTGCGCTCAGGCTCCCAAAAGTGTGGGATTTACAGGCGTGAGCGCTGC 420  
 QY 421 CTGGCTTTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG 480  
 Db 421 CTGGCTTTTATTTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG 480  
 QY 481 TGCAGTGTGTGATCACAGCTCAGTGCAGGCTTCAACTCTCTGAGATCAAGCATCTCTCTG 540  
 Db 481 TGCAGTGTGTGATCACAGCTCAGTGCAGGCTTCAACTCTCTGAGATCAAGCATCTCTCTG 540  
 QY 541 CCTCAGCTCTCCCAAGTAGCTGGGACCAAGACATGCAACACTACCTGGCTAAATTTTAA 600  
 Db 541 CCTCAGCTCTCCCAAGTAGCTGGGACCAAGACATGCAACACTACCTGGCTAAATTTTAA 600  
 QY 601 TTTTATTTTAAATTTTGTGAGACAGAGTCTCAACTCTGTCAACCAGGCTGGAGTGCAGT 660  
 Db 601 TTTTATTTTAAATTTTGTGAGACAGAGTCTCAACTCTGTCAACCAGGCTGGAGTGCAGT 660  
 QY 661 GGGCAATCTGGGCTCAGTGCNAACCTCTGCTCCCGGGTTCAAGTATATCTCTGCGCCCA 720  
 Db 661 GGGCAATCTGGGCTCAGTGCNAACCTCTGCTCCCGGGTTCAAGTATATCTCTGCGCCCA 720  
 QY 721 GCCTCTGAGTAGCTGGGACTACAGGCGCCACACAGCTAGCTAAATTTTGTATTTT 780  
 Db 721 GCCTCTGAGTAGCTGGGACTACAGGCGCCACACAGCTAGCTAAATTTTGTATTTT 780  
 QY 781 TAGTAGAGATGGGTTTCAACATTTGCGCAGGTTGATCTTGATCTCTGACCTTTGTGATC 840  
 Db 781 TAGTAGAGATGGGTTTCAACATTTGCGCAGGTTGATCTTGATCTCTGACCTTTGTGATC 840  
 QY 841 TGCCTGCTCCCGCTCCCAAGTGTGGGATTTACAGGCGTGAGCCACCAACCGCGCTTAA 900  
 Db 841 TGCCTGCTCCCGCTCCCAAGTGTGGGATTTACAGGCGTGAGCCACCAACCGCGCTTAA 900  
 QY 901 TTTTAAATTTTGTGTTGTAATGGAATCTCACTCTGTACCCAGGCTGGAGTGCAT 960  
 Db 901 TTTTAAATTTTGTGTTGTAATGGAATCTCACTCTGTACCCAGGCTGGAGTGCAT 960  
 QY 961 GGGCAATCTCGGCTCAGTGCNAACCTCTGCTCCCGGGTTCAAGCGATTTCTCTGTCTCA 1020  
 Db 961 GGGCAATCTCGGCTCAGTGCNAACCTCTGCTCCCGGGTTCAAGCGATTTCTCTGTCTCA 1020  
 QY 1021 GCCTCCCAAGCAGCTGGGATTTACGGGCACTGCGCACCAACCGCGCTAAATTTTGTATTT 1080  
 Db 1021 GCCTCCCAAGCAGCTGGGATTTACGGGCACTGCGCACCAACCGCGCTAAATTTTGTATTT 1080  
 QY 1081 TCATTAGAGCGGGGTTTCAACATTTGTCAGGCTGGTCTCAAACTCTCTGACCTCAGGT 1140  
 Db 1081 TCATTAGAGCGGGGTTTCAACATTTGTCAGGCTGGTCTCAAACTCTCTGACCTCAGGT 1140  
 QY 1141 GACCACCTGCTCAGGCTTCCAAAGTGTGGGATTTACAGGCGTGAGCCACCTCACCAG 1200  
 Db 1141 GACCACCTGCTCAGGCTTCCAAAGTGTGGGATTTACAGGCGTGAGCCACCTCACCAG 1200  
 QY 1201 CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGCTTTGCTATGTGCGCCAGGCT 1260  
 Db 1201 CCGGCTAAATTTAGATAAAAAATATGTAGCAATGGGGGCTTTGCTATGTGCGCCAGGCT 1260  
 QY 1261 GGTCTCAAACTTCTGCTTCAATGCAATCTTCAAAATGAGCCAAACACCCAGTCAC 1320  
 Db 1261 GGTCTCAAACTTCTGCTTCAATGCAATCTTCAAAATGAGCCAAACACCCAGTCAC 1320  
 QY 1321 CATTTTAAACAGTTACATCTTTATTTAGTATATACTAGAAAGTAATACAAATATGT 1380

Db 1321 CATTTTAAACAGTTACATCTTTATTTTAGTATACAGAAAGTAATACAATAAATCATGT 1380  
QY 1381 CAAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATATACTTTTAAACAAAGCTTTAGAG 1440  
Db 1381 CAAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTATATACTTTTAAACAAAGCTTTAGAG 1440  
QY 1441 CA 1442  
Db 1441 CA 1442

RESULT 12

US-10-910-173-1  
; Sequence 1, Application US/10910173  
; Publication No. US20050090441A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R  
; TITLE OF INVENTION: Inhibition of Neurodegeneration  
; FILE REFERENCE: 21486-047  
; CURRENT APPLICATION NUMBER: US/10/910,173  
; PRIOR FILING DATE: 2004-08-02  
; PRIOR APPLICATION NUMBER: US/09/872,968  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-910-173-1

Query Match 100.0%; Score 1442; DB 21; Length 1442;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTATTTTTCAGATGGAGTTTTCGCTCTGTGTCACAGGCTGGAGTGCATGGCGCAA 60  
Db 1 TTTTATTTTTCAGATGGAGTTTTCGCTCTGTGTCACAGGCTGGAGTGCATGGCGCAA 60  
QY 61 TCTCAGCTCACGCAACTCCGCTCCGCGGTTTCAAGGATTTCTCTGCTCAGCTCCC 120  
Db 61 TCTCAGCTCACGCAACTCCGCTCCGCGGTTTCAAGGATTTCTCTGCTCAGCTCCC 120  
QY 121 CAGTAGCTGGGATTAACGCAATGCAACCGCTCGGCTAATTTTGTATTTTATTTAG 180  
Db 121 CAGTAGCTGGGATTAACGCAATGCAACCGCTCGGCTAATTTTGTATTTTATTTAG 180  
QY 181 TAGAGATGGAGTTTCTCCATGTTGGTTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240  
Db 181 TAGAGATGGAGTTTCTCCATGTTGGTTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240  
QY 241 CTTCCGCTCCGCGCTCCCAAGTGTAGATACAGACTGGCCACCATGCGCGCTCTGCC 300  
Db 241 CTTCCGCTCCGCGCTCCCAAGTGTAGATACAGACTGGCCACCATGCGCGCTCTGCC 300  
QY 301 TGGCTAAATTTTGTGGTAGAACAAGGTTTCACTGATGTCGCAAGCTGCTCTCTGAGC 360  
Db 301 TGGCTAAATTTTGTGGTAGAACAAGGTTTCACTGATGTCGCAAGCTGCTCTCTGAGC 360  
QY 361 TCAAGCAGTCCACCTGCGCTCAGCCTCCAAAGTGTGGGATTAACAGGCGTGCAGCGTGC 420  
Db 361 TCAAGCAGTCCACCTGCGCTCAGCCTCCAAAGTGTGGGATTAACAGGCGTGCAGCGTGC 420  
QY 421 CTGGCTTTTATTTTATTTTATTTTAAACAGAGTGTCCCACTCTTAACCAAGTGAAG 480  
Db 421 CTGGCTTTTATTTTATTTTATTTTAAAGACACAGGTTGCCACTCTTACCCAGGATGAAG 480  
QY 481 TGCAGTGGTGTATCACAGCTCAGCTGAGCCTTCACTCTGAGATCAAGCATCTCTCTG 540  
Db 481 TGCAGTGGTGTATCACAGCTCAGCTGAGCCTTCACTCTGAGATCAAGCATCTCTCTG 540  
QY 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGGCTAATTTTAA 600

RESULT 13

US-09-964-666-4  
; Sequence 4, Application US/09964666  
; Patent No. US20020104108A1  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; Wands, Jack R.  
; TITLE OF INVENTION: Transgenic Animals and Cell Lines for  
; Screening Drugs Effective for the Treatment or Prevention

Db 541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGCACCATACCTTGGCTTAATTTTAA 600  
QY 601 TTTTATTTTAAATTTTGTGAGACAGAGTCTCAACTCTGTCAACAGGCTGGAGTGCAGT 660  
Db 601 TTTTATTTTAAATTTTGTGAGACAGAGTCTCAACTCTGTCAACAGGCTGGAGTGCAGT 660  
QY 661 GGGCGAATCTTGGCTCAGTGCACCTCTGCGCTCCGGGTTCAAGTTATTTCTCTGCCCA 720  
Db 661 GGGCGAATCTTGGCTCAGTGCACCTCTGCGCTCCGGGTTCAAGTTATTTCTCTGCCCA 720  
QY 721 GCCTCTCAGTAGCTGGGACTACAGGCGCCCAACAGCCTAGCTAATTTTGTATTTT 780  
Db 721 GCCTCTCAGTAGCTGGGACTACAGGCGCCCAACAGCCTAGCTAATTTTGTATTTT 780  
QY 781 TAGTAGAGATGGGTTTCAACATGTTTCGACAGGTTGATCTTGATCTCTGGACCTTGTGATC 840  
Db 781 TAGTAGAGATGGGTTTCAACATGTTTCGACAGGTTGATCTTGATCTCTGGACCTTGTGATC 840  
QY 841 TGCCTGCGCTCGGCTCCCAAGTGTGGGATTAACAGGCGTGCAGCACCAACCGCGGCTTA 900  
Db 841 TGCCTGCGCTCGGCTCCCAAGTGTGGGATTAACAGGCGTGCAGCACCAACCGCGGCTTA 900  
QY 901 TTTTATTTTGTGTTTGAATGGAATCTCACTCTGTACCCAGGCTGGAGTGCAT 960  
Db 901 TTTTATTTTGTGTTTGAATGGAATCTCACTCTGTACCCAGGCTGGAGTGCAT 960  
QY 961 GGCCTGCTCGGCTCAGTGCACCTCTGCTCCGGGCTCAAGGATTTCTCTGTCTCA 1020  
Db 961 GGCCTGCTCGGCTCAGTGCACCTCTGCTCCGGGCTCAAGGATTTCTCTGTCTCA 1020  
QY 1021 GCCTCCCAAGCAGCTGGGATTAACGGGCACTGCACCAACCGCGCTAATTTTGTATTT 1080  
Db 1021 GCCTCCCAAGCAGCTGGGATTAACGGGCACTGCACCAACCGCGCTAATTTTGTATTT 1080  
QY 1081 TCATTTAGAGGCGGGTTTCAACATATTTGTGAGGCTGGTCTCAAACTCTGACCTCAGGT 1140  
Db 1081 TCATTTAGAGGCGGGTTTCAACATATTTGTGAGGCTGGTCTCAAACTCTGACCTCAGGT 1140  
QY 1141 GACCCACCTGCTCAGCCTTCCAAAGTGTGGGATTAACAGGCTGAGCACTCAACCCAG 1200  
Db 1141 GACCCACCTGCTCAGCCTTCCAAAGTGTGGGATTAACAGGCTGAGCACTCAACCCAG 1200  
QY 1201 CCGGCTAATTTAGATAAAAAAATATGTAGCAATGGGGGCTCTGTGATTTGCCAGGCT 1260  
Db 1201 CCGGCTAATTTAGATAAAAAAATATGTAGCAATGGGGGCTCTGTGATTTGCCAGGCT 1260  
QY 1261 GGTCTCAAACTTCTGGCTTCATGCAATCCTTCCAAATGAGCCACCAACCCAGCAGTCA 1320  
Db 1261 GGTCTCAAACTTCTGGCTTCATGCAATCCTTCCAAATGAGCCACCAACCCAGCAGTCA 1320  
QY 1321 CATTTTAAACAGTTACATCTTTTATTTAGTATATCTAGAAAGTAAATCAATAAATGT 1380  
Db 1321 CATTTTAAACAGTTACATCTTTTATTTAGTATATCTAGAAAGTAAATCAATAAATGT 1380  
QY 1381 CAAACCTGCAAAATTCAGTAGTAAACAGAGTCTTTTATATACTTTTAAACAAAGCTTTAG 1440  
Db 1381 CAAACCTGCAAAATTCAGTAGTAAACAGAGTCTTTTATATACTTTTAAACAAAGCTTTAG 1440  
QY 1441 CA 1442  
Db 1441 CA 1442

of Alzheimer's Disease  
 NUMBER OF SEQUENCES: 14  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.  
 STREET: 1100 New York Ave., Suite 600  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20005-3934

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/964,666  
 FILING DATE: 28-Sep-2001  
 CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:  
 NAME: Emond, Robert W.  
 REGISTRATION NUMBER: 32,893  
 REFERENCE/DOCKET NUMBER: 0609.4370000

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-371-2600  
 TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1418 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: both

MOLECULE TYPE: cDNA  
 SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-964-666-4

Query Match 84.8%; Score 1223.4; DB 9; Length 1418;  
 Best Local Similarity 96.9%; Pred. No. 0;  
 Matches 1375; Conservative 0; Mismatches 31; Indels 13; Gaps 12;

QY 2 TTTTCTTTTGTAGTGGAGTTTCTGCTGTTGTTGCTCCAGGCTGGAGTGCATGGCGCAAT 61  
 DB 1 TTTTCTTTTGTAGTGGAGTTTCTGCTGTTGTTGCTCCAGGCTGGAGTGCATGGCGCAAT 60

QY 62 CTCAGCTCACCGCAACCTCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGCCTCCCC 121  
 DB 61 CTCAGCTCACCGCAACCTCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGCCTCCCC 120

QY 122 AGTA-GCTGGGATTACAGCATGTGCAACCGCTGGCTAATTTTGTATTTTGTAG 180  
 DB 121 AGTAGGCTGGGATTACAGCATGTGCA-CCAGCTCGGCTAATTTTGTATTTTGTAG 179

QY 181 TAGAGATGGAGTTTCTCCATGTTGTTGCTAGGCTGTTCTGAACTCCCGACCTCAGATGATC 240  
 DB 180 TAGAGATGGAGTTTCTCCATGTTGTTGCTAGGCTGTTCTGAACT-CCGACCTCAGATGATC 238

QY 241 CTCCTGCTCTCGGCTCTCCAAAGTGTAGATACAGGACTGGCCACCATGCGCGG-CTCTGC 299  
 DB 239 CTCCTGCTCTCGGCTCTCCAAAGTGTAGATACAGGACTGGCCACCATGCGGCTCTGC 298

QY 300 CTGGCTAATTTTGTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGTTCTCTGAG 359  
 DB 299 CTGGCTAATTTTGTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGTTCTCTGAG 358

QY 360 CTCAGCAGTCCACCTGCTCAGCTCCCAAGTGTGGATTACAGGCTGGCGCTG 419  
 DB 359 CTCAGCAGTCCACCTGCTCAGCTCCCAAGTGTGGATTACAGGCTGGCGCTG 418

QY 420 CTGGCTCTTTTATTTTATTTTATTTTAAAGACACAGGTTGCCACCTCTTACCAGGATGAA 479  
 DB 419 CTGGCTCTTTTATTTTATTTTATTTTAAAGACACAGGTTGCCACCTCTTACCAGGATGAA 478

QY 480 GTGAGTGGTGTGATCAAGCTCACTGAGGCTTCACTCTCTGAGATCAAGCATCTCTCT 539

DB 479 GTGAGTGGTGTGATCAAGCTCACTGAGCCTTCAACT-CTGAGATCAAGCATCTCTCT 537  
 QY 540 GCTCAGCTCC-AGTAGCTGGGACCAAGACATGACCACTACACTGCTGCTAATTTT 598  
 DB 538 GCTCAGCTCCAAAGTAGCTGGGACCAAGACATGACCACTACACTGCTGCTAATTTT 597  
 QY 599 TATTTTATTTTAAATTTTGTGAGACAGAGTCTCAACTCTGTCAACCCAGGCTGGAGTGA 658  
 DB 598 TATTTTATTTTAAATTTTGTGAGACAGAGTCTCAACTCTGTCAACCCAGGCTGGAGTGA 657  
 QY 659 GTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCCGGGTTCAAGTTATTTCTCTGCC 718  
 DB 658 GTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCCGGGTTCAAGTTATTTCTCTGCC 717  
 QY 719 CAGCTCTCTGAGTAGCTGGGACTACAGCGCCCAACAGCTAGCTAAATTTTCTTAT 778  
 DB 718 CAGCTCTCTGAGTAGCTGGGACTACAGCGCCCAACAGCTAGCTAAATTTTCTTAT 777  
 QY 779 TTTAGTAGAGTGGG-TTCAACATGTTGCGCAGGTTGAT-CTTGATCTCTGAGCCTTGT 836  
 DB 778 TTTAGTAGAGTGGGTTTCAACATGTTGCGCAGGTTGATCTGATCTCTGAGCCTTGT 837  
 QY 837 GATCTGCTGCTCGGCTCCCAAGTGTGCGGATTACAGG-CGTGAGCCACACGCCCG 895  
 DB 838 GATCTGCTGCTCGGCTCCCAAGTGTGCGGATTACAGGACGTGACGCCCCACGCCCG 897  
 QY 896 GCTATTTTAAATTTTGTGTTGTTGAAATGGAATCTCACTCTGTTTACCGAGCTGGAGT 955  
 DB 898 GCTATTTTAAATTTTGTGTTGTTGAAATGGAATCTCACTCTGTTTACCGAGCTGGAGT 957  
 QY 956 GCAATGGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGGCTCAAGCGATTCTCTCTG 1015  
 DB 958 GCAATGGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGGCTCAAGCGATTCTCTCTG 1017  
 QY 1016 TCTCAGCTCCCAAGCAGCTGGGATTACGGGCACTGCGCACCAACCCGCTAAATTTTGG 1075  
 DB 1018 TCTCAGCTCCCAAGCAGCTGGGATTACGGGCACTG-CACCACACCCCGCTAAATTTTGG 1076  
 QY 1076 TATTTTCAATAGAGCGGGGTTTCAACATATTTGTGAGGCTGCTCAAACTCTGACCT 1135  
 DB 1077 TATTTTCAATAGAGCGGGGTTTCAACATATTTGTGAGGCTGCTCAAACTCTGACCT 1136  
 QY 1136 CAGTGACCCACCTGCTCAGCTTCCAAAGTGTGGGATTACAGGCGTAGCCACCTCA 1195  
 DB 1137 CAGTGACCCACCTGCTCAGCTTCCAAAGTGTGGGATTACAGGCGTGA--CGCTCA 1194  
 QY 1196 CCGAGCGGCTAATTTAGATATAAATAATATGAGCAATGGGGGCTTCTGCTATGTTGCC 1255  
 DB 1195 CCGAGCGGCTAATTTAGATATAAATAATATGAGCAATGGGGGCTTCTGCTATGTTGCC 1254  
 QY 1256 AGGCTGCTCAAACTCTGCTCATGCAATCTTCCAAATGAGCCACACCCAGCC 1315  
 DB 1255 AGGCTGCTCAAACTCTGCTCATGCAATCTTCCAAATGAGCCACACCCAGCC 1314  
 QY 1316 AGTCACATTTTAAACAGTGTACATCTTTATTTTAGTATATAGAAAGTAAATCAATAA 1375  
 DB 1315 AGTCACA-TTTTAAACAGTGTACATCTTTATTTAGTATATAGAAAGTATACGATAAC 1373  
 QY 1376 CATGTCAACCTGCAATTCAGTAGTAACAGATTTCTTT 1414  
 DB 1374 ATGGCGGAACCTGCAATTCGAGTAGTACAGATCTTTT 1412

RESULT 14  
 US-09-964-412-4  
 ; Sequence 4, Application US/09964412  
 ; Patent No. US20020129391A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: de la Monte, Suzanne  
 ; ; Wands, Jack R.  
 ; TITLE OF INVENTION: Transgenic Animals and Cell Lines for  
 ; Screening Drugs Effective for the Treatment or Prevention  
 ; of Alzheimer's Disease

NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.  
STREET: 1100 New York Ave., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3934  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/964,412  
FILING DATE: 28-Sep-2001  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0609.4370000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1418 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-964-412-4  
Query Match 84.8%; Score 1223.4; DB 9; Length 1418;  
Best Local Similarity 96.9%; Pred. No. 0;  
Matches 1375; Conservative 0; Mismatches 31; Indels 13; Gaps 12;  
QY 2 TTTTCTTTTGGAGTGGAGTTTCTGCTCTGTTGCTCCAGGCTGGAGTGGCAATGGCGCAAT 61  
Db 1 TTTTCTTTTGGAGTGGAGTTTCTGCTCTGTTGCTCCAGGCTGGAGTGGCAATGGCGCAAT 60  
QY 62 CTCAGCTCACCGCAACCTCCGCTCCCGGTTTCAAGCGATTCTCTGCTCAGCTCAGCTCCGC 121  
Db 61 CTCAGCTCACCGCAACCTCCGCTCCCGGTTTCAAGCGATTCTCTGCTCAGCTCAGCTCCGC 120  
QY 122 AGTA-GCTGGGATTACAGGCATGTGCACCGCTCGGCTAATTTTGTATTTTTTTTAG 180  
Db 121 AGTAGGCTGGGATTACAGGCATGTGCA-CCAGCTCGGCTAATTTTGTATTTTTTTTAG 179  
QY 181 TAGAGATGGAGTTTCTCAATGTTGTTCAAGGCTGGTCTCGAATCTCCGACCTCAGATGATC 240  
Db 180 TAGAGATGGAGTTTCTCAATGTTGTTCAAGGCTGGTCTCGAATCTCCGACCTCAGATGATC 238  
QY 241 CCTCGCTCCGCTCCCGCTCCCGGTTTCAAGGCTGGTCTCGAATCTCCGACCTCAGATGATC 299  
Db 239 CTCCGCTCCGCTCCCGCTCCCGGTTTCAAGGCTGGTCTCGAATCTCCGACCTCAGATGATC 298  
QY 300 CTGGCTAATTTTGTGGTAGAAACAGGTTTCACTGATGTGCCAAGCTGGTCTCTCGAG 359  
Db 299 CTGGCTAATTTTGTGGTAGAAACAGGTTTCACTGATGTGCCAAGCTGGTCTCTCGAG 358  
QY 360 CTCAAGAGTGTCACTGCTCAGCTCCCAAGTGTGGGATTACAGGCTGGAGCGGTG 419  
Db 359 CTCAAGAGTGTCACTGCTCAGCTCCCAAGTGTGGGATTACAGGCTGGAGCGGTG 418  
QY 420 CCTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 479  
Db 419 CCTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 478  
QY 480 GTGCAGTGGTGTGATCACAGTCTCACTGAGCGCTTCAACTCTCTGAGATCAAGCATCTCTCT 539  
Db 479 GTGCAGTGGTGTGATCACAGTCTCACTGAGCGCTTCAACTCTCTGAGATCAAGCATCTCTCT 537

QY 540 GCCTCAGCCTCCC-AAGTAGCTGGGACCAAGAGATGACACCTACACTGCTGGCTAATTTT 598  
Db 538 GCCTCAGCCTCCC-AAAGTAGCTGGGACCAAGAGATGACACCTACACTGCTGGCTAATTTT 597  
QY 599 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 658  
Db 598 TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 657  
QY 659 GTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGTTTCAAGTATTATTTCTCTGCCC 718  
Db 658 GTGGCGCAATCTTGGCTCACTGCAACCTCTGCTCCCGGTTTCAAGTATTATTTCTCTGCCC 717  
QY 719 CAGCCTCTCTGAGTAGCTGGGACTACAGCGGCCACACAGCCTAGCTAATTTTGTATT 778  
Db 718 CAGCCTCTCTGAGTAGCTGGGACTACAGCGGCCACACAGCCTAGCTAATTTTGTATT 777  
QY 779 TTTAGTAGAGATGGGG-TTCAACCATGTTCCGACAGTTGAT-CTTGATCTCTGACCTTCT 836  
Db 778 TTTAGTAGAGATGGGGTTTCAACCATGTTCCGACAGTTGATGCTAGATCTCTTGACCTTCT 837  
QY 837 GATCTGCTGCTCGGCTCCCAAGAGTCTTGGGATTACAGG-CGTGAGCCACACGCGCCG 895  
Db 838 GATCTGCTGCTCGGCTCCCAAGAGTCTTGGGATTACAGGAGCTGACGCGCCACGCGCCG 897  
QY 896 GCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 955  
Db 898 GCTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 957  
QY 956 GCAATGGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGGCTCAAGCGATTCTCTCTG 1015  
Db 958 GCAATGGCCAAATCTCGGCTCACTGCAACCTCTGCTCCCGGCTCAAGCGATTCTCTCTG 1017  
QY 1016 TCTCAGCCTCCCAAGCAGCTGGGATTACGGGCACTGCGCACCTGCGCACCTGCTAATTTT 1075  
Db 1018 TCTCAGCCTCCCAAGCAGCTGGGATTACGGGCACTG-CACACACCGCGCTAATTTT 1076  
QY 1076 TATTTTCAATAGAGCGGGGTTTCACTATATTTTGTGAGCTGGTCTCAAACTCTCTGACT 1135  
Db 1077 TATTTTCAATAGAGCGGGGTTTCACTATATTTTGTGAGCTGGTCTCAAACTCTCTGACT 1136  
QY 1136 CAGGTGACCCACCTGCGCTCAGCTTCCAAAGTGTGGGATTACAGGCTGAGCCACCTCA 1195  
Db 1137 CAGGTGACCCACCTGCGCTCAGCTTCCAAAGTGTGGGATTACAGGCGTGA--CGCTCA 1194  
QY 1196 CCCAGCGGCTAATTTTAGATAAAAAATATGTAGCAATGGGGGGTCTTGTCTATGTTGCC 1255  
Db 1195 CCCAGCGGCTAATTTTAGATAAAAAATATGTAGCAATGGGGGGTCTTGTCTATGTTGCC 1254  
QY 1256 AGGCTGGTCTCAAACTTCTGGCTTCATGCAATCTCTTCCAAATGAGCCCAACACCCAGCC 1315  
Db 1255 AGGCTGGTCTCAAACTTCTGGCTTCATGCAATCTCTTCCAAATGAGCCCAACACCCAGCC 1314  
QY 1316 AGTCACATTTTAAACAGATTACATCTTTATTTTAGTACTAGAAAGTAAATACATAAA 1375  
Db 1315 AGTCACA-TTTTAAACAGATTACATCTTTATTTTAGTACTAGAAAGTAAATACATAAA 1373  
QY 1376 CATGTCAAACTCCAAATTCAGTAGTAACAGAGTTCTTT 1414  
Db 1374 ATGGCGAACCTGCBAATTCGAGTAGTACAGAGTCTTT 1412

RESULT 15  
US-09-964-667-4  
; Sequence 4, Application US/09964667  
; Publication No. US20030033621A1  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; Wands, Jack R.  
; TITLE OF INVENTION: Transgenic Animals and Cell Lines for  
; Screening Drugs Effective for the Treatment or Prevention  
; of Alzheimer's Disease  
; NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.  
 STREET: 1100 New York Ave., Suite 600  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20005-3934

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/964,667  
 FILING DATE: 28-Sep-2001  
 CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:  
 NAME: Esmond, Robert W.  
 REGISTRATION NUMBER: 32,893  
 REFERENCE/DOCKET NUMBER: 0609,4370000  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-371-2600  
 TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1418 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: both  
 MOLECULE TYPE: cDNA  
 SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
 US-09-964-667-4

Query Match 84.8%; Score 1223.4; DB 10; Length 1418;  
 Best Local Similarity 96.9%; Pred. No. 0;  
 Matches 1375; Conservative 0; Mismatches 31; Indels 13; Gaps 12;

QY	540	GCCTCAGCTCC-AAGTAGCTGGACCAAAACATGACACCTACACCTGGCTAATTTT	598
DB	538	GCCTCAGCTCCCAAGTAGCTGGACCAAAACATGACACCTACACCTGGCTAATTTT	597
QY	599	TATTTTATTTTAAATTTTGGAGACAGAGTCTCAACTCTGTCAACCCAGGCTGGAGTGA	658
DB	598	TATTTTATTTTAAATTTTGGAGACAGAGTCTCAACTCTGTCAACCCAGGCTGGAGTGA	657
QY	659	GTGGGCGCAATCTGGCTCACTGCAACCTCTGGCTCCCGGTTCAAGTTATTTCTCTGCC	718
DB	658	GTGGGCGCAATCTGGCTCACTGCAACCTCTGGCTCCCGGTTCAAGTTATTTCTCTGCC	717
QY	719	CAGCCTCTCCTAGTAGCTGGGACTACAGGCGCCACACCGCTAGCTAAATTTTCTAT	778
DB	718	CAGCCTCTCCTAGTAGCTGGGACTACAGGCGCCACACCGCTAGCTAAATTTTCTAT	777
QY	779	TTTAGTAGAGATGGGG-TTCACCATGTTGCGCAGGTTGAT-CTTGATCTCTGACCTTGT	836
DB	778	TTTAGTAGAGATGGGGTTTCAACATGTTGCGCAGGTTGATGCTAGATCTCTTGACCTTGT	837
QY	837	GATCTGCGCTCGCTCGGCTCACTGCAACCTCTGCGCTCCCGGCTCAAGCGATTTCTCTG	895
DB	838	GATCTGCGCTCGCTCGGCTCCCAAAAGTCTGGGATTTACAGGACGTGACGCCACCGCCG	897
QY	896	GCTTATTTTAAATTTTGTCTTTCGAAATGGAATCTCACTCTGTTTACCCAGGCTGGAGT	955
DB	898	GCCTATTTTAAATTTTGTCTTTCGAAATGGAATCTCACTCTGTTTACCCAGGCTGGAGT	957
QY	956	GCAATGGGCAAAATCTCGGCTCACTGCAACCTCTGCGCTCCCGGCTCAAGCGATTTCTCTG	1015
DB	958	GCAATGGGCAAAATCTCGGCTCACTGCAACCTCTGCGCTCCCGGCTCAAGCGATTTCTCTG	1017
QY	1016	TCTCAGCCTCCCAAGCAGCTGGGATTTACGGGCACTGCGGCACTCCACCAACCCCGCTAA	1075
DB	1018	TCTCAGCCTCCCAAGCAGCTGGGATTTACGGGCACTG-CACCACACCCCGCTAAATTTT	1076
QY	1076	TATTTTCAATTAGAGCGGGGTTTCAACATATTTTGTGAGGCTGGTCTCAAACTCCTGACCT	1135
DB	1077	TATTTTCAATTAGAGCGGGGTTTCAACATATTTTGTGAGGCTGGTCTCAAACTCCTGACCT	1136
QY	1136	CAGGTGACCCACCTCGCTCAGCCTTCCAAAGTCTGGGATTTACAGGCGTGA--CGCCTCA	1195
DB	1137	CAGGTGACCCACCTCGCTCAGCCTTCCAAAGTCTGGGATTTACAGGCGTGA--CGCCTCA	1194
QY	1196	CCAGCGGCTAAATTTAGATAAAATATATGCAATGGGGGCTTGTCTATGTTGCC	1255
DB	1195	CCAGCGGCTAAATTTAGATAAAATATATGCAATGGGGGCTTGTCTATGTTGCC	1254
QY	1256	AGGCTGGTCTCAAACTTCTGGCTTCAATGCAATCTTCCAAATGAGCCACAAACCCAGCC	1315
DB	1255	AGGCTGGTCTCAAACTTCTGGCTTCAATGCAATCTTCCAAATGAGCCACAAACCCAGCC	1314
QY	1316	AGTCAATTTTAAACAGTTACATCTTTATTTTATGATATATCTAGAAAGTAATCAATAAA	1375
DB	1315	AGTCAATTTTAAACAGTTACATCTTTATTTAGTATATCTAGAAAGTAATCAATAAA	1373
QY	1376	CATGTCAAACTGCAAAATTCAGTAGTAACAGAGTTCTTT	1414
DB	1374	ATGGCGGAACCTGCAAAATTCAGTAGTAACAGAGTTCTTT	1412

Search completed: September 15, 2005, 22:16:25  
 Job time : 1054 secs

QY	2	TTTTTTTTTGTAGATGGAGTTTTCGCTCTGTTGCCAGGCTGGAGTGAATGGCGCAAT	61
DB	1	TTTTTTTTTGTAGATGGAGTTTTCGCTCTGTTGCCAGGCTGGAGTGAATGGCGCAAT	60
QY	62	CTCAGCTCACCGCAACTCCGCTCCCGGTTCAAGCGATTTCTCTGCTCAGCCTCCCG	121
DB	61	CTCAGCTCACCGCAACTCCGCTCCCGGTTCAAGCGATTTCTCTGCTCAGCCTCCCG	120
QY	122	AGTA-GCTGGATACAGGATGTGACCCAGCTCGGCTAAATTTGTAATTTTGTAG	180
DB	121	AGTAGGCTGGGATTTACAGGATGTGCA-CCACGCTCGGCTAAATTTGTAATTTTGTAG	179
QY	181	TAGAGATGAGTTTCTCCATGTTGTCAGGCTGGTCTCGAATCTCCGACCTCAGATGATC	240
DB	180	TAGAGATGAGTTTCTCCATGTTGTCAGGCTGGTCTCGAATCTCCGACCTCAGATGATC	238
QY	241	CCTCGCTCTCGGCTCCCAAGTGTAGATACAGGACTGGCCACCAATGCCGCG-CTCTGC	299
DB	239	CTCGGCTCTCGGCTCCCAAGTGTAGATACAGGACTGGCCACCAATGCCGCGCTCTGC	298
QY	300	CTGGCTAAATTTTGGTGAAGAACAGGTTTCACTGATGTGCCAAGCTGGTCTCTGAG	359
DB	299	CTGGCTAAATTTTGGTGAAGAACAGGTTTCACTGATGTGCCAAGCTGGTCTCTGAG	358
QY	360	CTCAGCAGTCCACCTGCTCAGCCTCCCAAGTGTGGGATTTACAGGCGTGCAGCCGTG	419
DB	359	CTCAGCAGTCCACCTGCTCAGCCTCCCAAGTGTGGGATTTACAGGCGTGCAGCCGTG	418
QY	420	CCTGGCTTTTATTTTATTTTAAACAGAGTGTCCCACTCTTACCCAGGATGAA	479
DB	419	CCTGGCTTTTATTTTATTTTAAACAGAGTGTCCCACTCTTACCCAGGATGAA	478
QY	480	GTGAGTGTGTGATCACAGCTCACTGAGGCTTCAACTCTCTGAGATCAAGCATCTCTCT	539
DB	479	GTGAGTGTGTGATCACAGCTCACTGAGGCTTCAACT-CTGAGATCAAGCATCTCTCTCT	537

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2005, 19:28:46 ; Search time 18 Seconds  
(without alignments)  
2004.516 Million cell updates/sec

Title: US-09-380-203-2

Perfect score: 2034

Sequence: 1 MEFSLLPRLECNCAISRAH.....FIRGGVSPYLSGWSQTPDLR 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	382.5	18.8	673	F40201	artifact-warning s
2	315.5	15.5	613	C40201	artifact-warning s
3	315.5	15.5	627	A40201	artifact-warning s
4	291.5	14.3	597	E40201	artifact-warning s
5	197.5	9.7	579	D40201	artifact-warning s
6	175	8.6	499	S65657	alpha-1C-adrenergi
7	167	8.2	46	I54375	gene NF2 protein -
8	161	7.9	841	I78885	serine/threonine-s
9	157	7.7	39	I54374	gene NF2 protein -
10	151.5	7.4	301	A40201	artifact-warning s
11	139.5	6.9	100	A46010	X-linked retinopat
12	137.5	6.8	79	A56194	thromboxane A-2 re
13	137.5	6.8	407	T02670	probable thromboxa
14	128	6.3	53	A42442	integrin beta-1 ch
15	112	5.5	440	A26359	decay-accelerating
16	106.5	5.2	1125	T19193	hypothetical prote
17	105	5.2	522	T08711	gamma-adaptin homo
18	102.5	5.0	331	S59501	interferon recepto
19	99.5	4.9	331	A54295	interferon alpha/b
20	96.5	4.7	1643	1 RRGNV	genome polypeptin
21	96	4.7	458	E82175	conserved hypothet
22	92	4.5	964	T15342	hypothetical prote
23	91.5	4.5	542	S39608	transcription fact
24	91.5	4.5	631	A36749	genome polypeptin
25	91.5	4.5	1776	1 RRPYM	middle surface ant
26	90.5	4.4	282	1 SAVL64	DNA-directed DNA p
27	90.5	4.4	832	1 S20752	large surface anti
28	90	4.4	431	1 SAVLC2	probable xylogluca
29	89	4.4	299	2 F84785	

30	89	4.4	333	2 AI2131	hypothetical prote
31	89	4.4	1162	2 B97852	hypothetical prote
32	88	4.3	209	2 JC4244	heat-shock 27K pro
33	88	4.3	726	2 S18208	rabphilin-3A-inter
34	88	4.3	832	2 S71785	DNA-directed DNA p
35	87.5	4.3	494	2 S39607	transcription fact
36	87	4.3	226	1 JQ1574	major surface anti
37	87	4.3	226	2 JQ1574	surface antigen -
38	87	4.3	470	2 T05258	glycine hydroxymet
39	87	4.3	1153	2 A49676	nitric-oxide synth
40	86.5	4.3	360	2 T45956	hypothetical prote
41	86.5	4.3	1886	2 S04921	nuclear pore prote
42	86	4.2	262	2 AF2291	hypothetical prote
43	86	4.2	625	2 E96721	hypothetical prote
44	85.5	4.2	832	1 JQVLA1	DNA-directed DNA p
45	84.5	4.2	191	2 C72455	hypothetical prote

## ALIGNMENTS

### RESULT 1

F40201  
artifact-warning sequence (translated ALU class F) - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1992 #sequence\_revision 11-Aug-1995 #text\_change 19-May-2000  
C:Accession: F40201  
R:Claverie, J.M.  
personal communication, 1992  
A:Reference number: A40201  
A:Accession: F40201  
A:Molecule type: DNA  
A:Residues: 1-673 <CLA>  
R:Claverie, J.M.  
Genomics 12, 838-841, 1992  
A:Title: Identifying coding exons by similarity search: Alu-derived and other potential  
A:Reference number: A40200; MUID:92241891; PMID:1572661  
A:Contents: annotation  
C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames of  
in-frame stop codons are shown as 'X'.  
C:Comment: Any significant similarity of a predicted protein sequence to a portion of t

Query Match	18.8%;	Score 382.5;	DB 4;	Length 673;
Best Local Similarity	33.2%;	Pred No. 6.4e-26;		
Matches 128;	Conservative 19;	Mismatches 111;	Indels 127;	Gaps 17;
QY	16	ISAHNRLLPGSSDSDSPASVPAGITGMCTHARLILYFVLVEMEFHVQAGLELPTSD	75	
Db	371	VLACHNLGLGSSDSDSPASVSRVAGITGMGRHSWLI-YVFLIETQPHVDQAGLKLLTSSD	429	
QY	76	PSVSASQASRYRTGTHARLCLANFCGRNRVSLMCPSSPE-LKQSTCLSLPKCWYRRRA	134	
Db	430	-----LFSWSPKVLGXQAXATTPSXXXF---V 453		
QY	135	VPGLFILFLL-----RHRCP-----TLTQDEVQWCDHSHSLQSTPEIKHPASA	178	
Db	454	FGGFFFFFALFLRXALALTPELECGKFWLTAASTSWV-QAILLLSPVXLGLQAWA	512	
QY	179	SQVAGT-----KDMHHYTWLIFIFNFRLQSLNSVTQAGVQWRNLGSLQPLPPGFKLFSC	234	
Db	513	A-IPGXFWFYXXRHSFTMLI-----RLVSNXPQV-----ICPPG-----	546	
QY	235	PSLSSWDYRR-PPRLANFFVLVEMGFTMFARLILISGPCDLPASASQSGAGITGVSHA	293	
Db	547	--LPKCWDRREPHPAPAXXILFL-----GVFFF-----	572	
QY	294	RLIFNFCLEFESHVSVTQAGVQW-----PNLGLSLQPLPPGKRFSCLSLPSWMDYGLPHHP	350	
Db	573	--FFLLCFCFXDRPLLXHPG--WSAVAFGSLQPOPGFKRFSCLSLPCSWDYRAGPPFL	628	
QY	351	ANFCIFIRGGVSPYLSGWSQTPDLR	375	



```

Db      629 ANLCIFNRDVTSPCKSGMSQTPDLK 653

RESULT 2
C40201
artifact-warning sequence (translated ALU class C) - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000
C:Accession: C40201
R:Claverie, J.M.
personal communication, 1992
A:Reference number: A40201
A:Accession: C40201
A:Molecule type: DNA
A:Residues: 1-613 <CLA>
R:Claverie, J.M.
Genomics 12, 838-841, 1992
A:Title: Identifying coding exons by similarity search: Alu-derived and other potential
A:Reference number: A40200; MUID:92241891; PMID:1572661
A:Contents: annotation
C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames of c
in-frame stop codons are shown as 'X'.
C:Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match      15.5%; Score 315.5; DB 4; Length 613;
Best Local Similarity 33.4%; Pred. No. 5e-20;
Matches 98; Conservative 9; Mismatches 73; Indels 113; Gaps 8;

QY      6 LLPRLECNCAISAHNRNLRLPGSSDSPASASPVAGITGCTHARLILYFLVEMEFHLVQ 65
DB      424 LSPRLCEGGEISAHCKLHLPLGCHSPASAFVCGTGTGARTMPS-XFFVFLVEMGFCHVQ 482
QY      66 AGLELPTSDPVSASQSARYRTGHHARLCLANFCGRNRVSLMCPSPKQSTCLSLP 125
DB      483 AGLELLAS-----XSTHLCGLP 498
QY      126 KCWDYRR-----NAVGLFILFLRHRCPTLTODEVQWCHDSSLQSTPEIKHPPASA--- 178
DB      499 KCWDYRRLEPLHAPAXXFFFF-----XDRVSLC-----HPGWSAMAR 536
QY      179 SQVAGTKDMHYTTLWIFIFIFNFLRQSLNSVTQAGVQWRNLGSLQPLPPGFKLFSCPSLL 238
DB      537 SRLTASSRSRVAILL-----PQSAAYVGLQAPA-- 565
QY      239 SSWDYRRPRLANFFVFLVEMGTFMFARLILISGPCDLPASASQASAGITGVSH 291
DB      566 -----PCPANFLYFXKKMGFAMLARLVNSWPHDPTFPASQASAGITGVSH 610

RESULT 3
A40201
artifact-warning sequence (translated ALU class A) - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000
C:Accession: A40201
R:Claverie, J.M.
personal communication, 1992
A:Reference number: A40201
A:Accession: A40201
A:Molecule type: DNA
A:Residues: 1-627 <CLA>
R:Claverie, J.M.
Genomics 12, 838-841, 1992
A:Title: Identifying coding exons by similarity search: Alu-derived and other potential
A:Reference number: A40200; MUID:92241891; PMID:1572661
A:Contents: annotation
C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames of c
in-frame stop codons are shown as 'X'.
C:Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match      15.5%; Score 315.5; DB 4; Length 627;

```

```

Best Local Similarity 33.4%; Pred. No. 5.2e-20;
Matches 98; Conservative 13; Mismatches 67; Indels 115; Gaps 7;

QY      6 LLPRLECNCAISAHNRNLRLPGSSDSPASASPVAGITGCTHARLILYFLVEMEFHLVQ 65
DB      436 LSPRLCEGGEISAHCKLHLPGSRHSPPASASQVAGTTGARTTPG-XFFVFLVETGHRGSQ 494
QY      66 AGLELPTSDPVSASQSARYRTGHHARLCLANFCGRNRVSLMCPSPKQSTCLSLP 125
DB      495 DGLDLTL-----SXSAK-----LGLP 510
QY      126 KCWDYRRAAVGLFILFLRHRCPTLTODEVQWCHDSSLQSTPEIKHPPASASQVAGTK 185
DB      511 KCWDYRRRETAPG-----
QY      186 DMHYTTLWIFIFIF-----NFLRQSLNSVTO-----AGVQWRNLGSLQPLPPGFKLFSCPS 236
DB      523 ---XXXFLFFVIFIFRDGVSICRCQWSAVASRLTASSASRVHAILLPQPPKXGLQAPA 579
QY      237 LLSSWDYRRPRLANFFVFLVEMGTFMFARLILISGPCDLPASASQASAGITGV 289
DB      580 L-----RPAFLYFXRRRGFTVVARMVVISXPRDPPALASQASAGITGV 622

RESULT 4
E40201
artifact-warning sequence (translated ALU class E) - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1992 #sequence_revision 11-Aug-1995 #text_change 19-May-2000
C:Accession: E40201
R:Claverie, J.M.
personal communication, 1992
A:Reference number: A40201
A:Accession: E40201
A:Molecule type: DNA
A:Residues: 1-597 <CLA>
R:Claverie, J.M.
Genomics 12, 838-841, 1992
A:Title: Identifying coding exons by similarity search: Alu-derived and other potential
A:Reference number: A40200; MUID:92241891; PMID:1572661
A:Contents: annotation
C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames of c
in-frame stop codons are shown as 'X'.
C:Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match      14.3%; Score 291.5; DB 4; Length 597;
Best Local Similarity 30.5%; Pred. No. 6.6e-18;
Matches 118; Conservative 24; Mismatches 106; Indels 139; Gaps 20;

QY      5 LLLPRLECNCAISAHNRNLRLPGSSD-SPASASPVAGITGCTHARLILYF-----FLVEM 58
DB      310 LLLPRPEYDGTISP--QTLPGFKQFSPVASASVAGITGMRHHAQLILYFXRWGFSMLV 367
QY      59 EFLHVQAGLELPTSDPVSASQSARYT-----GHHARLC 95
DB      368 KLVNSQPQVIRPALNSQSAGIT-GMSYHTWXXFFFFFETFCSCCGRSTMAQSHRKL 426
QY      96 LANFCGRNRVSLMCPSPKQSTCLSLPCWDYRRAAVGLFILFLRHRCPTLTODE 155
DB      427 ---FLGSSNLLSQP---PEX-----LGLQAC-----ATTPSXVIF-----SRDG 461
QY      156 V-----QWCHDSSLQSTPEIKHPPASASQVAGTKDMHYTTLWIFIFIFNLRQSLNSVT 210
DB      462 VSPCWSSWSQTPNLRXSAPPXT-PKALGLQAXATTP-----GXXXFFFLRRSFALVA 512
QY      211 QAGVQWRNLGSLQPLPPGF-----KLFCPSLLSSWDYRR--PRLANFFVFLVEMGTFM 264
DB      513 QAGVRHNLTA-----NFAWVQAILSCSLPSSWDYRHAPPAPAN-FILVEMGFL-- 563
QY      265 ARLILISGFCDLPASASQASAGITGVSHHARLIFNFCFLFEMESHVSVTQAGVQWPNLGLSQP 324
DB      564 -----HVGQAGLKLPTSGD--- 577

```

QY 325 LPPGLKRFSCLSLPSSWDYGHLPHPA 351  
DB 578 -PPRL- -LPPKWDYRHEPLH 597

RESULT 5  
D40201  
Artifact-warning sequence (translated ALU class D) - human  
C:Species: Homo sapiens (man)  
C>Date: 31-Mar-1992 #sequence\_revision 11-Aug-1995 #text\_change 19-May-2000  
C:Accession: D40201  
R:Claverie, J.M.  
personal communication, 1992  
A:Reference number: A40201  
A:Accession: D40201  
A:Molecule type: DNA  
A:Residues: 1-579 <CLA>  
R:Claverie, J.M.  
Genomics 12, 838-841, 1992  
A:Title: Identifying coding exons by similarity search: Alu-derived and other potentially  
A:Reference number: A40200; MUID:92241891; PMID:1572661  
A:Contents: annotation  
C:Comment: This "warning" entry is a conceptual translation in all 6 reading frames of c  
in-frame stop codons are shown as 'x'.  
C:Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match 9.7%; Score 197.5; DB 4; Length 579;  
Best Local Similarity 28.1%; Pred. No. 1.4e-09;  
Matches 81; Conservative 11; Mismatches 69; Indels 127; Gaps 10;

QY 12 CN-GAISAHRNLRPGSSDSPASAPVAGITGCTHARLILYFFLVEMEFLHVGQAGLEL 70  
DB 404 CNHGSLS- -LDPPGQVILLPPPEXELGLQAYATRSXGLYLFVFEVGRHVAQVLEL 459

QY 71 PTSDPVSASQSAARYRTGTHARLCLANFCGRNRSVLMCPMSWPELKQSTCLSLPKCWDY 130  
DB 460 - - - - -LSWSIHLPW- - - - -LLKWDY 475

QY 131 RRAAV- - - - -PGFILFLRHRCPTLTQDEVQWCHSSLOFSPTPEIKHPASASQVAG 183  
DB 476 RRAAMLGXXKXFGQGLMPLREYRGAIMA- - - - -HXAL- - - - -TSQV- - - - -512

QY 184 TKDHHYTWLIFIFENFLRQSLNSVTQAGVQWRNLGSLQPLPPGKFLPSCPSLLSSWDY 243  
DB 513 - - - - -KXSSCPSLLSSWDY 526

QY 244 R-RPPRLANFFV-FLVEMGFTMFARLILSGPCDLPASASQAGITGV 289  
DB 527 RLMLPDLANFCIFFLXRWGFDMPLRLFLNSXAGAYTCHGSXAGITGV 574

RESULT 6  
S65657  
alpha-1C-adrenergic receptor splice form 2 - human  
N:Alternate names: alpha-1C-adrenoceptor isoform 2  
C:Species: Homo sapiens (man)  
C>Date: 22-Nov-1996 #sequence\_revision 22-Nov-1996 #text\_change 09-Jul-2004  
C:Accession: S65657; S65655  
R:Tanaka, T.  
submitted to the EMBL Data Library, July 1994  
A:Reference number: S65656  
A:Accession: S65657  
A:Molecule type: mRNA  
A:Residues: 1-499 <TAN>  
A:Cross-references: UNIPROT:Q13675; EMBL:D32202; NID:G927208; PIDN:BAA06901.1; PID:G9272  
R:Hirasawa, A.; Shibata, K.; Horie, K.; Takei, Y.; Ohika, K.; Tanaka, T.; Muramoto, N.;  
FEBS Lett. 363, 256-260, 1995  
A:Title: Cloning, functional expression and tissue distribution of human alpha (1C)-adren  
A:Reference number: S65654; MUID:95255557; PMID:7737411  
A:Accession: S65655  
A:Molecule type: mRNA

A:Residues: 424-499 <HIR>  
A:Cross-references: EMBL:D32202  
C:Genetics:  
A:Gene: GDB:ADRA1C; ADRA1L1  
A:Cross-references: GDB:128088; OMIM:104221  
A:Map position: 8p21-8p11.2  
C:Superfamily: vertebrate rhodopsin  
C:Keywords: alternative splicing; neurotransmitter receptor

Query Match 8.6%; Score 175; DB 2; Length 499;  
Best Local Similarity 58.0%; Pred. No. 1.1e-07;  
Matches 40; Conservative 5; Mismatches 20; Indels 4; Gaps 1;

QY 9 RLECNCAISAHRNLRPGSSDSPASAPVAGITGMC- - - - -THARLILYFFLVEMEFLHVG 64  
DB 431 RLECSGMILAHCNLRPGSRDSPASASQAAGTGDVPGRRHQAQLIFVLVETGPHVG 490

QY 65 QAGLELPTS 73  
DB 491 QDDLDLLTS 499

RESULT 7  
I54375  
gene NF2 protein - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 01-Nov-1996 #sequence\_revision 01-Nov-1996 #text\_change 16-Aug-2004  
C:Accession: I54375  
R:Arakawa, H.; Hayashi, N.; Nagase, H.; Ogawa, M.; Nakamura, Y.  
Hum. Mol. Genet. 3, 565-568, 1994  
A:Title: Alternative splicing of the NF2 gene and its mutation analysis of breast and c  
A:Reference number: I54375; MUID:94348501; PMID:8069299  
A:Accession: I54375  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-46 <RES>  
A:Cross-references: UNIPROT:Q16230; GB:S73853; NID:G688372; PIDN:AAB31736.1; PID:G68837  
C:Genetics:  
A:Gene: GDB:NF2  
A:Cross-references: GDB:120232; OMIM:101000  
A:Map position: 22q12.2-22q12.2  
C:Superfamily: laminin-type EGF-like homology

Query Match 8.2%; Score 167; DB 2; Length 46;  
Best Local Similarity 76.2%; Pred. No. 3.6e-08;  
Matches 32; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 302 FEMESHVTOAGVQVQWPNLGSLLQPLPPGLKRFSCLSLPSSWDY 343  
DB 4 FNCESCSVTLAGVQWRDLGLLQPLPPKRFKRFSCLSFPSSWDY 45

RESULT 8  
I78885  
serine/threonine-specific protein kinase (EC 2.7.1.1-) STK2 - human  
C:Species: Homo sapiens (man)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: I78885  
R:Lievadaku, E.N.; He, M.; Baptist, E.W.; Craven, R.J.; Cance, W.G.; Welcsh, P.L.; Simm  
Oncogene 9, 1977-1988, 1994  
A:Title: Two novel human serine/threonine kinases with homologies to the cell cycle reg  
A:Reference number: I58396; MUID:94268838; PMID:8208544  
A:Accession: I78885  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-841 <RES>  
A:Cross-references: UNIPROT:P51957; GB:L20321; NID:G348244; PIDN:AAA36658.1; PID:G34824  
C:Genetics:  
A:Gene: GDB:STK2  
A:Cross-references: GDB:374125  
A:Map position: 3p21.1-3p21.1  
C:Superfamily: human serine/threonine-specific protein kinase STK2; protein kinase homol  
C:Keywords: phosphotransferase

F;4-261/Domain: protein kinase homology <KIN>

Query Match 7.9%; Score 161; DB 1; Length 841;  
Best Local Similarity 46.2%; Pred. No. 3.6e-06;  
Matches 42; Conservative 14; Mismatches 25; Indels 10; Gaps 4;

QY 6 LLPRLECNCAISAHRNRLRFGSSDSPASASPVAGITGMCTHAR-LILYFFLVEMEFHVHG 64  
DB 460 LSPKLECSGTIIAHSNLRLLGSSDSPASASRVAGITGVCHHAQDQVAGECIIEKQ-----G 515

QY 65 QAGLELP---TSDDPSSVAS--QSARYRTGH 90  
DB 516 RIHPDLQPHNSGSEPSLSRQRKREQTEH 546

RESULT 9  
I54374  
gene NF2 protein - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 16-Aug-2004  
C:Accession: I54374  
R:Pykett, M.J.; Murphy, M.; Harnish, P.R.; George, D.L.  
Hum. Mol. Genet. 3, 559-564, 1994  
A:Title: The neurofibromatosis 2 (NF2) tumor suppressor gene encodes multiple alternative  
A:Reference number: I54374; MUID:94348500; PMID:8069298  
A:Accession: I54374  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-39 <RES>  
A:Cross-references: UNIPROT:Q16230; GB:L27065; NID:G463120; PIDN:AAA62422.1; PID:G463121  
C:Gene: GDB:NF2  
A:Gene: GDB:NF2  
A:Cross-references: GDB:120232; OMIM:101000  
A:Map position: 22q12.2-22q12.2  
C:Superfamily: laminin-type EGF-like homology

Query Match 7.7%; Score 157; DB 2; Length 39;  
Best Local Similarity 78.9%; Pred. No. 2.3e-07;  
Matches 30; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 306 SHSVTQAGVQWNLGSLQPLPGLKRFSCLSLPSWDY 343  
DB 1 SCSVTLAGVQWRDLGLQLPLPKPKFKFSCLSLPSWDY 38

RESULT 10  
B40201  
artifact-warning sequence (translated ALU class B) - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1992 #sequence\_revision 11-Aug-1995 #text\_change 19-May-2000  
C:Accession: B40201  
R:Claverie, J.M.  
personal communication, 1992  
A:Reference number: A40201  
A:Accession: B40201  
A:Molecule type: DNA  
A:Residues: 1-301 <CLA>  
R:Claverie, J.M.  
Genomics 12, 838-841, 1992  
A:Title: Identifying coding exons by similarity search: Alu-derived and other potential  
A:Reference number: A40200; MUID:92241891; PMID:1572661  
A:Contents: annotation  
C:Comment: this "warning" entry is a conceptual translation in all 6 reading frames of c  
in-frame stop codons are shown as 'X'.  
C:Comment: Any significant similarity of a predicted protein sequence to a portion of th

Query Match 7.4%; Score 151.5; DB 4; Length 301;  
Best Local Similarity 22.4%; Pred. No. 7.6e-06;  
Matches 66; Conservative 15; Mismatches 45; Indels 169; Gaps 10;

QY 6 LLPRLECNCAISAHRNRLRFGSSDSPASASPVAGITGMCT-----HARILLYFFLVEMEF 60

DB 158 MLPKLVNLSQAIIILQHP-----KVLGLQMLATMPCXXXTQLILFPF----- 211

QY 61 LHVGQAGLELPTSDPSSVASQSARYRTGHHARLCANFCGNNRVSLMCPSPWS--PELKQ 118  
DB 212 -----YRAG-----VLLCCPSWSXTPGLKQ 231

QY 119 STCLSLPKCWYRRAAVPGLFLFLFRLHRCPTLTQDEVQWCHSSLSQSTPIKHPASA 178  
DB 232 SSYFSIPKCWDYR-----C-----XFFCPA 251

QY 179 SOVAGTKDMHYTWLIFIFNFLQSLNSVTQAGVWRNLGSLQLPFGFKLFSCLPSLL 238  
DB 252 XXXPSX-----FYFIFELGSCYVAQAGLE----- 276

QY 239 SSWDYRRPRLANFFVFLVEMGFTWFARLLILISGPCDLPASASQAGITGVSHHA 293  
DB 277 -----LLVSSNP---PTSASQAGITDVSHHA 300

RESULT 11  
A46010  
X-linked retinopathy protein (C-terminal, clone XEH.8c) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A46010  
R:Wong, P.; Macdonald, I.M.; Sood, R.; Smith, C.; Pilon, R.; Tenniswood, M.  
Genomics 15, 467-471, 1993  
A:Title: Identification and partial characterization of a candidate gene for X-linked re  
A:Reference number: A46010; MUID:93224131; PMID:8468040  
A:Accession: A46010  
A>Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-100 <WON>  
A:Cross-references: UNIPROT:Q07826; GB:S58722; NID:G299470; PIDN:AAB26149.1; PID:G299471  
A>Note: sequence extracted from NCBI backbone (NCBIN:129339, NCBI:129340)

Query Match 6.9%; Score 139.5; DB 2; Length 100;  
Best Local Similarity 34.2%; Pred. No. 2.5e-05;  
Matches 38; Conservative 6; Mismatches 22; Indels 45; Gaps 2;

QY 197 FIFNFLQSLNSVTQAGVWRNLGSLQLPFGFKLFSCLPSLLSSWDYRRPRLANFFVFL 256  
DB 1 FFFPFETESCVAEAGVQWCDLGLSKSPPP----- 31

QY 257 VEMGFTMPARLILISGPCDLPASASQAGITGVSHARLIINFCLFEMESH 307  
DB 32 -----GSSDSPASASRVAGITGMHHHTQLIFVP-LVETGSH 66

RESULT 12  
A56194  
thromboxane A-2 receptor, endothelial - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 09-Jul-2004  
C:Accession: A56194  
R:Raychowdhury, M.K.; Yukawa, M.; Collins, L.J.; McGrail, S.H.; Kent, K.C.; Ware, J.A.  
J. Biol. Chem. 270, 7011, 1995  
A:Reference number: A56194; MUID:95204505; PMID:7896853  
A:Accession: A56194  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-79 <RAY>  
A:Cross-references: UNIPROT:Q9UCY2; GB:U11271  
C:Superfamily: prostaglandin E receptor EP1

Query Match 6.8%; Score 137.5; DB 2; Length 79;  
Best Local Similarity 49.3%; Pred. No. 2.8e-05;  
Matches 37; Conservative 8; Mismatches 25; Indels 5; Gaps 2;

QY 6 LLPRLECNCAISAHRNRLRFGSSDSPASASPVAGITGMCTHARLILYFFLVEMEFHVHG 65  
DB 5 LWPSEYSGTISAHNCNLRLLPGSSDSRASASRAAGITGV-SHCARPCMLFDPEFDLL---- 59

C;Species: Homo sapiens (man)  
C;Date: 05-Oct-1988 #sequence revision 05-Oct-1988 #text\_change 09-Jul-2004  
C;Accession: A26359; A39702; S16187; S23138; A27258  
R;Garas, I.W.; Davitz, M.A.; Rhee, L.; Weddell, G.; Martin Jr., D.W.; Nussenzweig, V.  
Nature 325, 545-549, 1987  
A;Title: Cloning of decay-accelerating factor suggests novel use of splicing to generate  
A;Reference number: A26359; MUID:87115845; PMID:2433596  
A;Accession: A26359  
A;Molecule type: mRNA  
A;Residues: 1-440 <CAR>  
A;Cross-references: UNIPROT:P08174; GB:M30142  
R;Ewulonu, U.K.; Ravi, L.; Medof, M.E.  
Proc. Natl. Acad. Sci. U.S.A. 88, 4675-4679, 1991  
A;Title: Characterization of the decay-accelerating factor gene promoter region.  
A;Reference number: A39702; MUID:91271256; PMID:1711208  
A;Accession: A39702  
A;Molecule type: DNA  
A;Residues: 1-79, T, 81-104 <EWU>  
A;Cross-references: GB:M64356  
A;Note: the authors translated the codon AGT for residue 85 as Met  
R;Nakano, Y.; Sugita, Y.; Ishikawa, Y.; Choi, N.H.; Tobe, T.; Tomita, M.  
Biochim. Biophys. Acta 1074, 326-330, 1991  
A;Title: Isolation of two forms of decay-accelerating factor (DAF) from human urine.  
A;Reference number: S16187; MUID:91291869; PMID:1712233  
A;Accession: S16187  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 35-47 <BIO>  
R;Nakano, Y.; Sumida, K.; Kikutani, N.; Miura, N.H.; Tobe, T.; Tomita, M.  
Biochim. Biophys. Acta 1116, 235-240, 1992  
A;Title: Complete determination of disulfide bonds localized within the short consensus  
A;Reference number: S23138; MUID:92305034; PMID:1377029  
A;Accession: S23138  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 35-41;65-68;79-81;93-103;128-134;143-145;155-159;162-168;188-192;203-204;21  
R;Sugita, Y.; Negoro, T.; Matsuda, T.; Sakamoto, T.; Tomita, M.  
J. Biochem. 100, 143-150, 1986  
A;Title: Improved method for the isolation and preliminary characterization of human DA  
A;Reference number: A27258; MUID:87008461; PMID:2428813  
A;Accession: A27258  
A;Molecule type: protein  
A;Residues: 35, 'X', 37, 'G', 39-51, 'P', 53-55, 'X', 57-58, 'X', 60-63 <SUG>  
A;Note: Gly-35 and Leu-38 were also found  
C;Comment: For an alternative splice form, see PIR:B26359  
C;Genetics:  
A;Gene: GDB:DAF  
A;Cross-references: GDB:119088; OMIM:125240  
A;Map position: lq32-lq32  
C;Superfamily: decay-accelerating factor; complement factor H repeat homology  
C;Keywords: alternative splicing; glycoprotein.  
F.1-34/Domain: signal sequence #status predicted <SIG>  
F.35-440/Product: decay-accelerating factor 1 #status predicted <MAT>  
F.36-94/Domain: complement factor H repeat homology <FH01>  
F.98-158/Domain: complement factor H repeat homology <FH02>  
F.163-220/Domain: complement factor H repeat homology <FH03>  
F.225-283/Domain: complement factor H repeat homology <FH04>  
F.95/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.5%; Score 112; DB 2; Length 440;  
Best Local Similarity 20.6%; Pred. No. 0.038;  
Matches 87; Conservative 43; Mismatches 100; Indels 192; Gaps 22;

QY 22 LRLGSSDSP--ASASPVAGITGMCTHARLLLYFLVEMEFHVQGAGLELTSDPPSVS 79  
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB 70 VKIPEKDSVICLKGQSWSDEECNR:-:-----CEVPTRLN---S 106  
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 80 ASOSARYRTGHARCLANFCGRNRVSLWC-----PSWSPKLQSTCLSLPKCWYRR 132  
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

DB 107 ASLKQPITYQN-----FPVGTVVECRGYRRFSLSPLK----TCLQLNK-W---- 151  
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY 133 AAAPGLFILFFLRHCPTLTDEVQCDHSSLQSPTEIKGPASASQVAGTKDMHHYT 192  
::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```
Db 152 -----STAVECKKSC-PNPGEIRN---CQIDVPG-----G 179
QY 193 LIFIFINF-----LRQSLNS---VTQAGVQWRNLGSLQPLPPGFKLFSCPSLLSSWD 242
Db 180 ILFGATISFSCNTGKLFQSTSSFCILSGSSVQWSD-----ELPECREIY-CPA----- 227
QY 243 YRRPRLANFFV-----FLVEMGFTMPARLILI-----SGP----- 273
Db 228 ---PFQIDNGIIOGERDHYGRQSVTYACNKGFTMIGEHSIYCTVNNDEGEWSGPPPECR 284
QY 274 -----CDLP-----ASASOSAGITGVSHH----- 292
Db 285 GKSLTSKVPPTVQKPTTVNVPTEVSPTSQTKTTTTPNAQATRSTPVSRTTKHFETT 344
QY 293 -----ARLIFNCLFEMESHVTVQAGVQWRNLGSLQPLPPGFKLFSCPSLLSSWD 341
Db 345 PNKGSOTTSGTTRLLSG-----SRPVTQAGMRWCDRSSLQSRTPGFKRSFHFSLPSSW 397
QY 342 DY 343
Db 398 YY 399
```

Search completed: September 13, 2005, 19:35:16  
Job time : 21 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2005, 19:23:06 ; Search time 72 Seconds  
(without alignments)  
2014.377 Million cell updates/sec

Title: US-09-380-203-2

Perfect score: 2034

Sequence: 1 MEFSLLRLRCNGAISAH.....PIRGVSPYSLGWSQTDLR 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_15Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2034	100.0	375	3 AAB01399	Aab01399 Neuron-as
2	2034	100.0	375	5 ABB81538	Abb81538 Neural th
3	2034	100.0	375	5 AAE29142	Aae29142 AD7c-NTP
4	2034	100.0	375	6 ABR63268	Abr63268 AD7c-NTP
5	2034	100.0	375	6 ADA84017	Ada84017 Human POM
6	2034	100.0	375	6 ABU03024	Abu03024 Human neu
7	2034	100.0	375	6 ABB99774	Abb99774 Amino aci
8	2034	100.0	375	6 ABJ19445	Abj19445 AD7C-neur
9	2034	100.0	375	7 ADB37642	Adb37642 Human neu
10	2034	100.0	375	8 ADR14409	Adr14409 Human NF-
11	1415.5	69.6	397	2 AAR95913	Aar95913 Neural th
12	928	45.6	381	4 AAU30235	Aau30235 Novel hum
13	927.5	45.6	382	4 AAU32610	Aau32610 Novel hum
14	927.5	45.6	382	4 AAU31818	Aau31818 Novel hum
15	927.5	45.6	382	4 AAU32707	Aau32707 Novel hum
16	849	41.7	324	4 AAU29573	Aau29573 Novel hum
17	849	41.7	324	4 AAU29579	Aau29579 Novel hum
18	822.5	40.4	411	4 ABG08428	Abg08428 Novel hum
19	739	36.3	317	4 AAU33200	Aau33200 Novel hum
20	735	36.1	290	4 ABG21913	Abg21913 Novel hum
21	734.5	36.1	361	5 ABG68738	Abg68738 Human.pro
22	729	35.8	449	5 ABJ01057	Abj01057 Ovary cel
23	701.5	34.5	332	5 ADK34410	Adk34410 Novel hum
24	609	29.9	257	4 AA010294	Aa010294 Human pol
25	590	29.0	276	4 ABG07919	Abg07919 Novel hum

## ALIGNMENTS

RESULT 1

AAB01399

ID AAB01399 standard; protein; 375 AA.

XX AC AAB01399;

XX DT 20-OCT-2000 (first entry)

XX DE Neuron-associated protein.

XX KW Neuron associated protein; NEUAP; neurological disorder; epilepsy;

XX KW ischemic cerebrovascular disease; stroke; cerebral neoplasm;

XX KW Alzheimer's disease; Pick's disease; Huntington's disease; dementia;

XX KW Parkinson's disease; demyelinating disease; meningitis; prion disease;

XX KW kuru; Creutzfeldt-Jakob disease; neurofibromatosis; cerebral palsy;

XX KW muscular dystrophy; central nervous system; CNS;

XX KW peripheral nervous system; PNS; myopathy; schizophrenia;

XX KW atrophic keratosis; arteriosclerosis; atherosclerosis; bursitis;

XX KW cirrhosis; hepatitis; mixed connective tissue disease; MCTD;

XX KW myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;

XX KW autoimmune disease; inflammation; acquired immunodeficiency syndrome;

XX KW AIDS; Addison's disease; adult respiratory distress syndrome; allergy;

XX KW ankylosing spondylitis; amyloidosis; anaemia; asthma;

XX KW Werner syndrome; trauma; human.

XX OS Homo sapiens.

XX WO200034477-A2.

XX 15-JUN-2000.

XX 10-DEC-1999; 99WO-US030408.

XX 11-DEC-1998; 98US-00210083.

XX 09-FEB-1999; 99US-0119365P.

XX 16-MAR-1999; 99US-0124687P.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Yue H, Baughn MR, Hillman JL, Lai P, Au-Young J;

XX Yang J, Lu DAM, Azimzai Y;

XX WPI; 2000-423423/36.

XX New human neuron-associated proteins and polynucleotides encoding them,  
XX useful for diagnosis, treatment and prevention of cell proliferative  
XX disorders including cancer, neuronal and neurological disorders.

Adk36936 Novel hum  
Aao06014 Human pol  
Abb11464 Human neu  
Ade09985 Novel pro  
Aau30686 Novel hum  
Aau31823 Novel hum  
Ade09441 Novel pro  
Aau33304 Novel hum  
Adk36978 Novel hum  
Aau31782 Novel hum  
Aau31513 Novel hum  
Abg07921 Novel hum  
Aau32615 Novel hum  
Aau30391 Novel hum  
Aau31857 Novel hum  
Aau31980 Novel hum  
Aau30455 Novel hum  
Aam25396 Human pro  
Aau31786 Novel hum  
Aau31834 Novel hum

26 575 28.3 302 5 ADK36936  
27 574 28.2 185 4 AAO06014  
28 569 28.0 241 4 ABB11464  
29 567.5 27.9 196 7 ADE09985  
30 563 27.7 202 4 AAU30686  
31 561 27.6 250 4 AAU31823  
32 561 27.6 250 7 ADE09441  
33 560 27.5 296 4 AAU33304  
34 554.5 27.3 232 5 ADK36978  
35 553 27.2 175 4 AAU31782  
36 550.5 27.1 215 4 AAU31513  
37 546 26.8 215 4 ABG07921  
38 543.5 26.7 216 4 AAU32615  
39 524 25.8 304 4 AAU30391  
40 522.5 25.7 175 4 AAU31857  
41 521.5 25.6 418 4 AAU31980  
42 520 25.6 396 4 AAU30455  
43 517 25.4 213 4 AAM25396  
44 504.5 24.8 196 4 AAU31786  
45 503 24.7 189 4 AAU31834

PS	Disclosure; Page 143-144; 145pp; English.	DE	Neural thread protein (NTP).
XX		XX	
CC	Human neuron-associated proteins (NEUP) can be used for for treating or	KW	Neural thread protein; NTP; Harlil peptide; Alzheimer's disease;
CC	preventing a disorder associated with decreased expression or activity of	KW	Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;
CC	NEUP. Antagonists of NEUP are useful for treating or preventing	XX	hypoxia; ischaemia; cerebral infarction.
CC	disorder associated with increased expression or activity of NEUP. NEUP	OS	Homo sapiens.
CC	or their fragments or derivatives are useful for treating neurological	XX	
CC	disorder such as epilepsy, ischemic cerebrovascular disease, stroke,	XX	WO200234915-A2.
CC	cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's	XX	
CC	disease, dementia and Parkinson's disease. NEUPs are also useful for	XX	
CC	treating other demyelinating diseases, bacterial and viral meningitis,	PD	02-MAY-2002.
CC	prion diseases including kuru, Creutzfeldt-Jakob disease, nutritional and	XX	
CC	metabolic diseases of the nervous system, neurofibromatosis, other	XX	
CC	developmental disorders of the central nervous system, cerebral palsy,	XX	25-OCT-2001; 2001WO-US042813.
CC	neurolept disorders, autonomic nervous system disorders, cranial	XX	
CC	nerve disorders, spinal cord diseases, muscular dystrophy and other	PR	27-OCT-2000; 2000US-00697590.
CC	neuromuscular disorders, peripheral nervous system disorders, inherited,	XX	
CC	metabolic, endocrine, and toxic myopathies, mental disorders including,	XX	(NYMO-) NYMOX PHARM CORP.
CC	mood, anxiety and schizophrenic disorders, a cell proliferative disorder	XX	
CC	such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis,	XX	
CC	cirrhosis, hepatitis, mixed connective tissue disease (MCTD),	XX	
CC	myelofibrosis, paroxysmal nocturnal haemoglobinuria, cancers of the	XX	
CC	adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, and an	XX	
CC	autoimmune/inflammatory disorder such as acquired immunodeficiency	XX	
CC	sndrome (AIDS), Addison's disease, adult respiratory distress syndrome,	XX	
CC	allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, Werner	XX	
CC	syndrome, complications of cancer, hemodialysis, and extracorporeal	XX	
CC	circulation, viral, bacterial, fungal parasitic, protozoal, and	XX	
CC	helminthic infections, and trauma. This protein was designated g3002527	XX	
XX		XX	
SQ	Sequence 375 AA;	XX	
	Query Match 100.0%; Score 2034; DB 3; Length 375;	CC	The present invention describes a neural thread protein (NTP) peptide
	Best Local Similarity 100.0%; Pred. No. 1e-198;	CC	having an amino acid sequence selected from ABB81511 to ABB81529 and
	Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	CC	their homologues, which are referred collectively as Harlil peptides (1).
		CC	(1) have neuroprotective, nootropic, vasotropic and cerebroprotective
QY	1 MEPSLLLPRLCNGAISAHNRLLPGSSDPSASPVAGITGMCTHARLILYFFLVEMEF 60	CC	activities, and can be used in peptide therapy. The Harlil peptide
DB	1 MEPSLLLPRLCNGAISAHNRLLPGSSDPSASPVAGITGMCTHARLILYFFLVEMEF 60	CC	sequences can be used as analogues for NTP in therapeutic or diagnostic
QY	61 LHVGAGLELPTSDPSVSASQARYRTGHHARLCLANFCGRRNRYSLMCPSPKQST 120	CC	assays by replacing NTP with the peptide in such an assay. The Harlil
DB	61 LHVGAGLELPTSDPSVSASQARYRTGHHARLCLANFCGRRNRYSLMCPSPKQST 120	CC	peptides are also useful as a trap material in a diagnostic or
QY	121 CLSLPKCWDYRAAIVPGLFILFHLRCPPTLTQDEVQWCHDSSLOPSTPEIKHPASASQ 180	CC	therapeutic assay. Therefore, the Harlil peptides are useful in binding
DB	121 CLSLPKCWDYRAAIVPGLFILFHLRCPPTLTQDEVQWCHDSSLOPSTPEIKHPASASQ 180	CC	assays, protein and antibody purification, therapeutics or diagnostics.
QY	181 VAGTKDMHHYTWLIFIFNLRQSLNSVTQAGVQWRNLGSLQPLPPGPKLFCPSLLSS 240	CC	In particular, the peptides are also useful for diagnosing Alzheimer's
DB	181 VAGTKDMHHYTWLIFIFNLRQSLNSVTQAGVQWRNLGSLQPLPPGPKLFCPSLLSS 240	CC	disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
QY	241 WDYRPPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300	CC	glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides
DB	241 WDYRPPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300	CC	are also useful as targets for drug development for the treatment of
QY	301 LFEMESHVSQTQGVQWPNLGSLOPLPPGLKRRFSCLSLPSWDYGHLPHPANFCIFIRGG 360	CC	these diseases. The present sequence represents a neural thread protein
DB	301 LFEMESHVSQTQGVQWPNLGSLOPLPPGLKRRFSCLSLPSWDYGHLPHPANFCIFIRGG 360	CC	given in the exemplification of the present invention
QY	361 VSPYLSGWSQTPDLR 375	XX	
DB	361 VSPYLSGWSQTPDLR 375	XX	
		SQ	Sequence 375 AA;
		QY	Query Match 100.0%; Score 2034; DB 5; Length 375;
		DB	Best Local Similarity 100.0%; Pred. No. 1e-198;
		QY	Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
		DB	Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
		QY	1 MEPSLLLPRLCNGAISAHNRLLPGSSDPSASPVAGITGMCTHARLILYFFLVEMEF 60
		DB	1 MEPSLLLPRLCNGAISAHNRLLPGSSDPSASPVAGITGMCTHARLILYFFLVEMEF 60
		QY	61 LHVGAGLELPTSDPSVSASQARYRTGHHARLCLANFCGRRNRYSLMCPSPKQST 120
		DB	61 LHVGAGLELPTSDPSVSASQARYRTGHHARLCLANFCGRRNRYSLMCPSPKQST 120
		QY	121 CLSLPKCWDYRAAIVPGLFILFHLRCPPTLTQDEVQWCHDSSLOPSTPEIKHPASASQ 180
		DB	121 CLSLPKCWDYRAAIVPGLFILFHLRCPPTLTQDEVQWCHDSSLOPSTPEIKHPASASQ 180
		QY	181 VAGTKDMHHYTWLIFIFNLRQSLNSVTQAGVQWRNLGSLQPLPPGPKLFCPSLLSS 240
		DB	181 VAGTKDMHHYTWLIFIFNLRQSLNSVTQAGVQWRNLGSLQPLPPGPKLFCPSLLSS 240
		QY	241 WDYRPPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300
		DB	241 WDYRPPRLANFFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300
		QY	301 LFEMESHVSQTQGVQWPNLGSLOPLPPGLKRRFSCLSLPSWDYGHLPHPANFCIFIRGG 360
		DB	301 LFEMESHVSQTQGVQWPNLGSLOPLPPGLKRRFSCLSLPSWDYGHLPHPANFCIFIRGG 360
		QY	361 VSPYLSGWSQTPDLR 375
		DB	361 VSPYLSGWSQTPDLR 375
		RESULT 2	
		ABB81538	
		ID ABB81538	standard; protein; 375 AA.
		XX	
		AC ABB81538;	
		XX	
		DT 02-SEP-2002 (first entry)	
		XX	



QY 301 LFMESHVTVQGVQWPNLGSQPLPPGLKRFSCSLPSSWDYGHLPHPANFCIFIRGG 360  
 DB |||||  
 QY 301 LFMESHVTVQGVQWPNLGSQPLPPGLKRFSCSLPSSWDYGHLPHPANFCIFIRGG 360  
 DB |||||  
 QY 361 VSPYLSGWSQTPDLR 375  
 DB |||||  
 QY 361 VSPYLSGWSQTPDLR 375  
 DB |||||

RESULT 3

AAE29142  
 ID AAE29142 standard; protein; 375 AA.

AC AAE29142;

DT 27-JAN-2003 (first entry)

DE AD7c-NTP protein.

KW Neural thread protein; NTP; hyperplasia; hypertrophy; arteriosclerosis;  
 KW haemorrhoid; gene therapy; tumour; vascular disease; atherosclerosis;  
 KW inflammatory disease; nutritional deficiency disease; genetic disease;  
 KW autoimmune disease; metabolic disease; traumatic disease; intoxication;  
 KW infectious disease; congenital malformation; enzyme deficiency disease;  
 KW amyloid disease; fibrosis disease; storage disease; radiation disease;  
 KW poisoning; environmental disease; endocrine disease; protein therapy;  
 KW degenerative disease; mechanical disease; AD7c-NTP protein.

OS Unidentified.

XX WO200274323-A2.

PN 26-SEP-2002.

PD 08-MAR-2002; 2002WO-1B001959.

PF 08-MAR-2001; 2001US-0273957P.

PR (AVER/) AVERBACK P.

PA AVERBACK P;

PI WPI; 2002-759864/82.

XX N-PSDB; AAD46671.

DR Treating a condition in a patient requiring removal or destruction of  
 PT cells, such as a benign or malignant tumor of a tissue or an inflammatory  
 PT disease, comprises administering a neural thread protein (NTP) or a NTP  
 PT gene to a mammal.

PS Example 2; Fig 1; 70pp; English.

XX The invention relates to a method for treating a condition in a patient  
 CC requiring removal or destruction of cells. The method involves  
 CC administering to a mammal a neural thread protein (NTP), or administering  
 CC to a tumour or other target cell a NTP gene, where the expression of the  
 CC NTP gene is induced resulting in expression of the NTP protein. The  
 CC method and NTP are useful for treating a condition in a patient requiring  
 CC removal or destruction of cells, such as a benign or malignant tumour of  
 CC a tissue, a hyperplasia, hypertrophy, or overgrowth of a tissue,  
 CC preferably tonsillar hypertrophy or prostatic hyperplasia, a virally,  
 CC bacterially, or parasitically altered tissue, or a malformation of a  
 CC tissue. Other conditions include a cosmetic modification to a tissue,  
 CC such as removal of unwanted facial hair, warts or unwanted fatty tissue,  
 CC a vascular disease, particularly atherosclerosis or arteriosclerosis,  
 CC haemorrhoids, or varicose veins, an inflammatory disease, autoimmune  
 CC disease, metabolic disease, hereditary/genetic disease, traumatic disease,  
 CC or physical injury, nutritional deficiency disease, infectious disease,  
 CC congenital malformation, amyloid disease, fibrosis disease, storage  
 CC disease, enzyme deficiency disease, poisoning, intoxication, degenerative  
 CC disease, radiation disease, environmental disease, endocrine disease or  
 CC mechanical disease. The invention is useful in protein therapy and gene  
 CC therapy. The present sequence is AD7c-NTP protein

XX Sequence 375 AA;  
 SQ Query Match  
 Best Local Similarity 100.0%; Score 2034; DB 5; Length 375;  
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MEFSLLLPRLECNCAISAHRLNRLPGSSDSDSPASASPVAGITGCTHARLILYFFLVEMEF 60  
 DB |||||  
 QY 1 MEFSLLLPRLECNCAISAHRLNRLPGSSDSDSPASASPVAGITGCTHARLILYFFLVEMEF 60  
 DB |||||  
 QY 61 LHVGQAGLELPTSDSDSPASASQARYRTGHARLCLANFCGNNRVSLMCPSPSPKQST 120  
 DB |||||  
 QY 61 LHVGQAGLELPTSDSDSPASASQARYRTGHARLCLANFCGNNRVSLMCPSPSPKQST 120  
 DB |||||  
 QY 121 CLSLPKCWDYRAAVPGLFILLFRLHRCPTLTQDSVQWCDHSSLQSPSTPEIKHPASASQ 180  
 DB |||||  
 QY 121 CLSLPKCWDYRAAVPGLFILLFRLHRCPTLTQDSVQWCDHSSLQSPSTPEIKHPASASQ 180  
 DB |||||  
 QY 181 VAGTKDMHHYTWLIFIFIFNLRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240  
 DB |||||  
 QY 181 VAGTKDMHHYTWLIFIFIFNLRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240  
 DB |||||  
 QY 241 WDYRPPRLANFFVFLVEMGTFMFARLILISGPCDLPASASQAGITGVSHARLIFNFC 300  
 DB |||||  
 QY 241 WDYRPPRLANFFVFLVEMGTFMFARLILISGPCDLPASASQAGITGVSHARLIFNFC 300  
 DB |||||  
 QY 301 LFMESHVTVQGVQWPNLGSQPLPPGLKRFSCSLPSSWDYGHLPHPANFCIFIRGG 360  
 DB |||||  
 QY 301 LFMESHVTVQGVQWPNLGSQPLPPGLKRFSCSLPSSWDYGHLPHPANFCIFIRGG 360  
 DB |||||  
 QY 361 VSPYLSGWSQTPDLR 375  
 DB |||||  
 QY 361 VSPYLSGWSQTPDLR 375  
 DB |||||

RESULT 4

ABR63268  
 ID ABR63268 standard; protein; 375 AA.

XX ABR63268;

DT 28-AUG-2003 (first entry)

DE AD7c-NTP protein.

KW Cytostatic; Antibacterial; Immunosuppressive; Antiinflammatory;  
 KW neural thread protein; NTP; tumour.

OS Unidentified.

XX WO2003008443-A2.

XX 30-JAN-2003.

XX 19-JUL-2002; 2002WO-CA001105.

XX 19-JUL-2001; 2001US-0306150P.

XX 19-JUL-2001; 2001US-0306161P.

XX 16-NOV-2001; 2001US-0331477P.

XX (NYMO-) NYMOX CORP.

XX AVERBACK PA;

XX WPI; 2003-247999/24.

XX Novel neural thread protein peptide, referred as cell death peptide,  
 XX useful for treating prostatic hyperplasia, psoriasis, eczema, dermatosis,  
 XX atherosclerosis, cosmetic modification to skin, throat, mouth, muscle.

PS Disclosure; Fig 1; 77pp; English.

CC The present invention relates to a neural thread protein (NTP) peptide  
 CC referred to as cell death peptide. Thought to be cytostatic,  
 CC antibacterial, immunosuppressive and antiinflammatory. It is useful for  
 CC treating a condition in a patient requiring removal or destruction of  
 CC cells, for treating a condition such as benign or malignant tumor,  
 CC inflammatory disease, autoimmune disease and infectious disease. The  
 CC peptide useful for treatment is derived from the amino acid sequence for  
 CC a pancreatic thread protein. The peptide is conjugated, linked or bound  
 CC to a molecule chosen from antibody or its fragment, antibody-like binding  
 CC molecule, where the molecule has a higher affinity for binding to a tumor  
 CC or other target than binding to other cells. Treatment using NTP peptides  
 CC can remove benign tumors with less risk and fewer of the undesirable side  
 CC effects of surgery. The present sequence is an NTP amino acid sequence  
 XX  
 SQ Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 6; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 1e-198;  
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLPLECNGAISAHNRNLRPGSSDPSASAPVAGITGCTHARLILYFFLVEMEF 60  
 DB 1 MEFSLLPLECNGAISAHNRNLRPGSSDPSASAPVAGITGCTHARLILYFFLVEMEF 60

QY 61 LHVQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRNRLVSLMCPSPBELKOST 120  
 DB 61 LHVQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRNRLVSLMCPSPBELKOST 120

QY 121 CLSLPKCWDYRAAVPGLFILFHLRHCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180  
 DB 121 CLSLPKCWDYRAAVPGLFILFHLRHCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180

QY 181 VAGTKDMHHYTWLIFIFNLRQSLNSVTQAGVQWRNLSLQPLPPGKLFSCPSLSS 240  
 DB 181 VAGTKDMHHYTWLIFIFNLRQSLNSVTQAGVQWRNLSLQPLPPGKLFSCPSLSS 240

QY 241 WDYRPPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300  
 DB 241 WDYRPPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300

QY 301 LFEMESHVTOAGVQWPNLGSLOPLPPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGG 360  
 DB 301 LFEMESHVTOAGVQWPNLGSLOPLPPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGG 360

QY 361 VSPYLSGWSQTPDLR 375  
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 5  
 ID ADA84017 standard; protein; 375 AA.  
 AC ADA84017;  
 XX  
 XX 20-NOV-2003 (first entry)  
 XX  
 XX Human POM80 protein.  
 XX  
 KW human; marker; expressed sequence tag; EST; arabidopsis; tumour;  
 KW stress-induced phenotype; hyperosmotic stress; colon cancer; immunogen;  
 KW vaccine.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2002103028-A2.  
 XX  
 XX 27-DEC-2002.  
 XX  
 XX 30-MAY-2002; 2002NO-IB004189.  
 XX  
 XX 30-MAY-2001; 2001US-0293999P.  
 PR 22-OCT-2001; 2001US-0330457P.  
 PR

PR 19-FEB-2002; 2002US-0357144P.  
 XX  
 PA (BIOM-) BIOMEDICAL CENT.  
 XX  
 XX Baranova AV, Yankovsky NK, Kozlov AP, Lobashev AV, Krukovskaya LL;  
 XX WPI; 2003-175241/17.  
 DR N-PSDB; ADA84016.  
 XX  
 PT Determining if a nucleic acid is a marker for a phenotype/cell type of  
 PT interest, by global comparison of expressed sequence tags known to be  
 PT expressed in the phenotype/cell type with all ESTs expressed in normal  
 PT tissue.  
 XX  
 XX Claim 29; Page 398-400; 516pp; English.  
 XX  
 CC The invention relates to a novel method for determining if a nucleic acid  
 CC is a marker for a predetermined phenotype/cell type of interest from a  
 CC biological species. The method comprises performing a global comparison  
 CC of a group of expressed sequence tags (ESTs) known to be expressed in the  
 CC phenotype/cell type of interest with all ESTs expressed in normal tissue  
 CC in order to identify ESTs that are preferentially expressed in the  
 CC phenotype/cell of interest. A method of the invention is useful for  
 CC determining whether a nucleic acid is a marker for a predetermined  
 CC phenotype or cell type of interest from a biological species, preferably  
 CC Arabidopsis or human. The cell type of interest is an abnormal cell such  
 CC as a tumour cell, and the predetermined phenotype is a stress-induced  
 CC phenotype such as hyperosmotic stress or high salt conditions. A method  
 CC of the invention is also useful for determining the progression of colon  
 CC cancer in a human, for detecting a tumour cell, and for regulating or  
 CC preventing the growth of a tumour cell. An antibody of the invention is  
 CC useful for detecting the absence or presence of peptides encoded by  
 CC tumour-associated markers. A polypeptide of the invention is useful as an  
 CC immunogen for vaccinating an animal. The present sequence represents a  
 CC tumour-associated antigen of the invention.  
 XX  
 SQ Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 6; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 1e-198;  
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLPLECNGAISAHNRNLRPGSSDPSASAPVAGITGCTHARLILYFFLVEMEF 60  
 DB 1 MEFSLLPLECNGAISAHNRNLRPGSSDPSASAPVAGITGCTHARLILYFFLVEMEF 60

QY 61 LHVQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRNRLVSLMCPSPBELKOST 120  
 DB 61 LHVQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRNRLVSLMCPSPBELKOST 120

QY 121 CLSLPKCWDYRAAVPGLFILFHLRHCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180  
 DB 121 CLSLPKCWDYRAAVPGLFILFHLRHCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180

QY 181 VAGTKDMHHYTWLIFIFNLRQSLNSVTQAGVQWRNLSLQPLPPGKLFSCPSLSS 240  
 DB 181 VAGTKDMHHYTWLIFIFNLRQSLNSVTQAGVQWRNLSLQPLPPGKLFSCPSLSS 240

QY 241 WDYRPPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300  
 DB 241 WDYRPPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300

QY 301 LFEMESHVTOAGVQWPNLGSLOPLPPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGG 360  
 DB 301 LFEMESHVTOAGVQWPNLGSLOPLPPGLKRFSCLSLPSWDYGHLPHPANFCIFIRGG 360

QY 361 VSPYLSGWSQTPDLR 375  
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 6  
 ABU03024

ID XX  
AC XX  
XX XX  
DT XX  
XX XX  
DE XX  
XX XX  
KW ABU03024 standard; protein; 375 AA.  
AC ABU03024;  
XX XX  
DT 20-JAN-2003 (first entry)  
XX XX  
DE Human neural thread protein AD7C-NTP.  
XX XX  
KW Neural thread protein; NTP-peptide; AD7C-NTP; surgical excision;  
KW transplantation; grafting; chemotherapy; immunotherapy; vaccination;  
KW ablation; cryotherapy; laser therapy; phototherapy; gene therapy;  
KW radiation; tumor; hyperplasia; hypertrophy; overgrowth of tissue;  
KW malformation of tissue; tonsillary hypertrophy; prostatic hyperplasia;  
KW cosmetic modification; vascular disease; atherosclerosis;  
KW arteriosclerosis; haemorrhoid; varicose vein; inflammatory disease;  
KW autoimmune disease; metabolic disease; traumatic disease;  
KW physical injury; nutritional deficiency disease; infectious disease;  
KW amyloid disease; fibrosis disease; storage disease;  
KW congenital malformation; enzyme deficiency disease; poisoning;  
KW intoxication; environmental disease; radiation disease;  
KW endocrine disease; degenerative disease; mechanical disease.  
XX XX  
OS Homo sapiens.  
XX XX  
FN WO200297030-A2.  
XX XX  
PD 05-DEC-2002.  
XX XX  
XX 24-MAY-2002; 2002WO-CA000759.  
XX XX  
XX 25-MAY-2001; 2001US-0293156P.  
XX XX  
PA (NYMO-) NYMOX CORP.  
XX XX  
PI Averbach PA;  
XX XX  
DR WPI; 2003-041406/03.  
XX XX  
PT Novel peptides similar in amino acid sequence to neural thread proteins  
PT (NTP), useful for treating unwanted cellular proliferations such as  
PT malignant tumors and prostatic hyperplasia.  
XX XX  
PS Disclosure; Fig 1; 78pp; English.  
XX XX  
CC The invention describes an NTP-peptide (I) comprising at least one amino  
CC acid sequence corresponding to part of the amino acid sequence of a  
CC neural thread protein, AD7C-NTP. The invention provides a method of  
CC treating a condition requiring removal or destruction of cells of a  
CC mammal comprising administering to a mammal, a therapeutic amount of (I).  
CC The treatment is administered to the mammal before, during or after  
CC surgical excision, transplantation, grafting, chemotherapy,  
CC immunotherapy, vaccination, thermal or electrical ablation, cryotherapy,  
CC laser therapy, phototherapy, gene therapy and/or radiation. The method is  
CC useful for treatment of benign or malignant tumour; hyperplasia,  
CC hypertrophy or overgrowth of tissue; virally, bacterially or  
CC parasitically altered tissue; malformation of tissue selected from lung,  
CC breast, stomach, pancreas, prostate, bladder, bone, ovary, skin, kidney,  
CC sinus, colon, intestine, rectum, esophagus, heart, spleen, salivary  
CC gland, blood, brain and its coverings, spinal cord, muscle, connective  
CC tissue, adrenal, parathyroid, thyroid, uterus, testis, pituitary,  
CC reproductive organs, liver, hair, gall bladder, eye, ear, nose, throat,  
CC tonsils, mouth and lymph nodes and lymphoid system; tonsillary  
CC hypertrophy; prostatic hyperplasia; cosmetic modification to a tissue;  
CC vascular disease (atherosclerosis or arteriosclerosis); haemorrhoids;  
CC varicose veins; inflammatory disease; autoimmune disease; metabolic  
CC disease; hereditary/genetic disease; traumatic disease; physical injury;  
CC nutritional deficiency disease; infectious disease; amyloid disease;  
CC fibrosis disease; storage disease; congenital malformation; enzyme  
CC deficiency disease; poisoning; intoxication; environmental disease;  
CC radiation disease; endocrine disease; degenerative disease and mechanical  
CC disease. This is the amino acid sequence of the human neural thread  
CC protein AD7C-NTP  
XX XX

SQ Sequence 375 AA;  
Query Match 100.0%; Score 2034; DB 6; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1e-198;  
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEFSLLPRLCNGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLYEMEF 60  
DB 1 MEFSLLPRLCNGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLYEMEF 60  
QY 61 LHVGQAGLELPSTDPSVSASQASARYTGHARLCLANFCGRNRVSLMCPSPKQST 120  
DB 61 LHVGQAGLELPSTDPSVSASQASARYTGHARLCLANFCGRNRVSLMCPSPKQST 120  
QY 121 CUSLPCXWDYRAA VPGILFELFELRHRCPTLTQDEVQWCDHSLQSPSTEIKHPPASQ 180  
DB 121 CUSLPCXWDYRAA VPGILFELFELRHRCPTLTQDEVQWCDHSLQSPSTEIKHPPASQ 180  
QY 181 VAGTKDMHHYTWLIFIFNFRLQSLNSVTQAGVQWRNLGSLQPLPPGFKLFCSPCLLS 240  
DB 181 VAGTKDMHHYTWLIFIFNFRLQSLNSVTQAGVQWRNLGSLQPLPPGFKLFCSPCLLS 240  
QY 241 WDYRPPRLANFPFVLEMGFTMFARLILISGFCPLPASASQASAGITGVSHARLIFNFC 300  
DB 241 WDYRPPRLANFPFVLEMGFTMFARLILISGFCPLPASASQASAGITGVSHARLIFNFC 300  
QY 301 LFEMESHVSVTQAGVQWRNLGSLQPLPPGFKLFCSPCLLSQPLPPGFKLFCSPCLLS 360  
DB 301 LFEMESHVSVTQAGVQWRNLGSLQPLPPGFKLFCSPCLLSQPLPPGFKLFCSPCLLS 360  
QY 361 VSPYLSGWSQTPDLR 375  
DB 361 VSPYLSGWSQTPDLR 375  
RESULT 7  
ABB99774  
ID ABB99774 standard; protein; 375 AA.  
AC ABB99774;  
XX XX  
DT 24-MAR-2003 (first entry)  
XX XX  
DE Amino acid sequence of human neuronal thread protein AD7C-NTP.  
XX XX  
KW Human; neuronal thread protein; AD7C-NTP; Alzheimer's disease; histone;  
KW neurodegeneration; in vivo gene expression; amphipathic compound;  
KW gene therapy.  
XX XX  
OS Homo sapiens.  
XX XX  
FN WO200299036-A2.  
XX XX  
PD 12-DEC-2002.  
XX XX  
PF 28-MAY-2002; 2002WO-US016429.  
XX XX  
PR 01-JUN-2001; 2001US-00872968.  
XX XX  
PA (RHOD-) RHODE ISLAND HOSPITAL.  
XX XX  
Wands JR, De La Monte SM;  
XX XX  
WPI; 2003-140605/13.  
XX XX  
DR N-P5DB; ABZ23236.  
XX XX  
PT Inducing prolonged in vivo gene expression in mammal by contacting  
PT neuronal tissue with composition comprising Alzheimer's disease-  
PT associated neural thread protein 7c antisense nucleic acid, histone,  
PT amphipathic compound.  
XX XX  
PS Disclosure; Page 35; 69pp; English.  
XX XX

CC The present sequence represents a human neuronal thread protein AD7c-NTP.  
 CC AD7c-NTP is overexpressed in brains with Alzheimer's disease at early and  
 CC intermediate stages of neurodegeneration. The expression of AD7c-NTP may  
 CC be reduced using the method of the invention. The specification describes  
 CC a method for inducing prolonged in vivo gene expression in a mammal. The  
 CC method comprises contacting a non-muscular tissue with a composition  
 CC comprising a nucleic acid, histone and an amphipathic compound. The  
 CC method is useful for inducing prolonged in vivo gene expression in non-  
 CC muscular tissue of a mammal, e.g. neuronal tissue, central nervous system  
 CC (CNS) tissue, tissue comprising a post-mitotic neuronal cell, cortical  
 CC neuronal cell or hippocampal neuronal cell, glial cell, or vascular  
 CC endothelial cell. The method is useful in gene therapy applications to  
 CC treat Alzheimer's disease, where the composition comprises antisense  
 CC AD7c-NTP nucleic acid

SQ Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 6; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 1e-198;  
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDPSASPVAGITGCTHARLILYFFLVEMEF 60  
 DB 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDPSASPVAGITGCTHARLILYFFLVEMEF 60

QY 61 LHVGQAGLELPTSDPSVSASQARYTGHARLCLANFCGRNRYSLMCPSPKQST 120  
 DB 61 LHVGQAGLELPTSDPSVSASQARYTGHARLCLANFCGRNRYSLMCPSPKQST 120

QY 121 CLSLPKCDWYRAAIVPGLFIFLFRHRCPTLTQDEVQVQWCHSSLOPSTPEIKHPPASASQ 180  
 DB 121 CLSLPKCDWYRAAIVPGLFIFLFRHRCPTLTQDEVQVQWCHSSLOPSTPEIKHPPASASQ 180

QY 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240  
 DB 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240

QY 241 WDYRRPRLANFPFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300  
 DB 241 WDYRRPRLANFPFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300

QY 301 LFEMESHVSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 360  
 DB 301 LFEMESHVSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 360

QY 361 VSPYLSGWSQTPDLR 375  
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 8

ABU19445  
 ID ABU19445 standard; protein; 375 AA.  
 XX AC ABU19445;  
 XX DT 27-MAR-2003 (first entry)  
 XX DE AD7c-neural thread protein.  
 XX KW Nootropic; neuroprotective; cell death; tissue necrosis; NTP;  
 KW neural thread protein; neurodegenerative disorder; Alzheimer's disease.  
 XX Unidentified.  
 OS WO200292115-A2.  
 PN 21-NOV-2002.  
 PD 16-MAY-2002; 2002WO-CA000712.  
 XX 16-MAY-2001; 2001US-0290971P.  
 XX

PA (NYMO-) NYMOX CORP.  
 XX Averbach PA;  
 XX WPI; 2003-129234/12.  
 DR N-PSDB; AAL54224.  
 XX Preventing and/or inhibiting cell death and/or tissue necrosis in a  
 PT tissue for treating a neurodegenerative disorder, e.g. Alzheimer's  
 PT disease, by contacting the live tissue with at least one segment of  
 PT neural thread proteins (NTP).  
 XX Disclosure; Fig 1; 60pp; English.  
 XX The invention relates to a novel method for preventing and/or inhibiting  
 CC cell death and/or tissue necrosis in a tissue comprising contacting the  
 CC live tissue with at least one segment of neural thread proteins (NTP).  
 CC The methods are composition are useful for treating a neurodegenerative  
 CC disorder, such as Alzheimer's disease. This sequence represents the AD7c-  
 CC NTP protein relating to the invention

SQ Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 6; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 1e-198;  
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDPSASPVAGITGCTHARLILYFFLVEMEF 60  
 DB 1 MEFSLLLPRLECNCAISAHNRNLRPGSSDPSASPVAGITGCTHARLILYFFLVEMEF 60

QY 61 LHVGQAGLELPTSDPSVSASQARYTGHARLCLANFCGRNRYSLMCPSPKQST 120  
 DB 61 LHVGQAGLELPTSDPSVSASQARYTGHARLCLANFCGRNRYSLMCPSPKQST 120

QY 121 CLSLPKCDWYRAAIVPGLFIFLFRHRCPTLTQDEVQVQWCHSSLOPSTPEIKHPPASASQ 180  
 DB 121 CLSLPKCDWYRAAIVPGLFIFLFRHRCPTLTQDEVQVQWCHSSLOPSTPEIKHPPASASQ 180

QY 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240  
 DB 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240

QY 241 WDYRRPRLANFPFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300  
 DB 241 WDYRRPRLANFPFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300

QY 301 LFEMESHVSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 360  
 DB 301 LFEMESHVSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 360

QY 361 VSPYLSGWSQTPDLR 375  
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 9

ADB37642  
 ID ADB37642 standard; protein; 375 AA.  
 XX AC ADB37642;  
 XX DT 04-DEC-2003 (first entry)  
 XX DE Human neural thread protein AD7c-NTP.  
 XX KW Cytostatic; Antitumour; Antipsoriatic; Dermatological;  
 KW Antiatherosclerotic; Antiarteriosclerotic; Vasotropic; Antiinflammatory;  
 KW Immunosuppressive; Tranquillizer; Antiemetic; Virucide; AD7c-NTP;  
 XX neural thread protein; neuritic sprouting.  
 OS Homo sapiens.  
 XX

PN WO2003008444-A2.  
 XX 30-JAN-2003.  
 XX 19-JUL-2002; 2002WO-CA001106.  
 XX 19-JUL-2001; 2001US-0306150P.  
 PR 19-JUL-2001; 2001US-0306150P.  
 PR 16-NOV-2001; 2001US-0331477P.  
 XX (NYMO-) NYMOX CORP.  
 PA  
 XX  
 XX Averbach PA, Gemmell J;  
 PI WPI; 2003-248000/24.  
 XX N-PSDB; ADB37519.  
 DR  
 XX Novel Related peptide or AD7c-neural thread peptide, useful for treating  
 PT unwanted cellular proliferations, glandular hyperplasia, unwanted facial  
 PT hair, warts and unwanted fatty tissue.  
 XX  
 XX Disclosure; Fig 1; 109pp; English.  
 PS  
 XX The present invention relates to AD7c-neural thread protein (NTP) and  
 CC related proteins and peptides (I; ADB37528-ADB37641). The sequences are  
 CC useful for treating a condition in a patient requiring removal or  
 CC destruction of cells. The condition can be selected from benign or  
 CC malignant tumour of a tissue, hyperplasia, hypertrophy or overgrowth of a  
 CC tissue, virally, bacterially or parasitically altered tissue, or  
 CC malformation of a tissue, where the tissue is selected from lung, breast,  
 CC stomach, pancreas, prostate, bladder, bone, ovary, skin, kidney, sinus,  
 CC colon, intestine, stomach, rectum, oesophagus, heart, spleen, salivary  
 CC gland, blood, brain and its coverings, spinal cord and its coverings,  
 CC muscle, connective tissue, adrenal, parathyroid, thyroid, uterus, testis,  
 CC pituitary, reproductive organs, liver, gall bladder, eye, ear, nose,  
 CC throat, tonsils, mouth, lymph nodes and lymphoid tissue. The condition is  
 CC preferably tonsillary hypertrophy, prostatic hyperplasia, psoriasis,  
 CC eczema, dermatosis, cosmetic modification to a tissue (skin, eye, ear,  
 CC nose, throat, mouth, muscle, connective, hair or breast tissue), vascular  
 CC disease (atherosclerosis or arteriosclerosis), haemorrhoids, varicose  
 CC veins, inflammatory disease, autoimmune disease, metabolic disease,  
 CC hereditary/genetic disease, traumatic disease or physical injury,  
 CC nutritional deficiency disease, infectious disease, amyloid disease,  
 CC fibrosis disease, storage disease, congenital malformation, enzyme  
 CC deficiency disease, poisoning, intoxication, environmental disease,  
 CC radiation disease, endocrine disease, degenerative disease and mechanical  
 CC disease. The peptides are useful for treating unwanted cellular  
 CC proliferations, glandular (e.g. prostate) hyperplasia, unwanted facial  
 CC hair, warts and unwanted fatty tissue, or for preparing antibodies that  
 CC recognize and/or bind to Related proteins, Related peptides or NTP  
 CC peptides. The present sequence is AD7c-NTP, used to illustrate the  
 CC invention. AD7c-NTP is a ~41kd membrane associated phosphoprotein with  
 CC functions associated with neuritic sprouting.  
 XX  
 XX Sequence 375 AA;  
 SQ  
 Query Match 100.0%; Score 2034; DB 7; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 1e-198;  
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MEFSLLLPLECNGAISAHNRLRLPGSSDPSASFPVAGITGCTHARLILYFFLVEMEF 60  
 DB 1 MEFSLLLPLECNGAISAHNRLRLPGSSDPSASFPVAGITGCTHARLILYFFLVEMEF 60  
 QY 61 LVHQAGLELPTSDPSVASQSARYRTGHARLCLANFCGNNRVSLMCPSPSELKQST 120  
 DB 61 LVHQAGLELPTSDPSVASQSARYRTGHARLCLANFCGNNRVSLMCPSPSELKQST 120  
 QY 121 CLSLPKCWDYRAAVPGLFPLFLRHRCPTLTQDEVQWCHSSLOPSTPEIKHPASASQ 180  
 DB 121 CLSLPKCWDYRAAVPGLFPLFLRHRCPTLTQDEVQWCHSSLOPSTPEIKHPASASQ 180  
 QY 181 VAGTKDMHHYTLWLIFFIFNLRQSLNSVTQAGVQWRNLGSLQPLPPGPKLFSPCSLLSS 240

Db 181 VAGTKDMHHYTLWLIFFIFNLRQSLNSVTQAGVQWRNLGSLQPLPPGPKLFSPCSLLSS 240  
 QY 241 WDYRRPPRLANFFVFLVEMGFTMFARLILISGFCDLPASASQAGITGVSHARLIFNFC 300  
 Db 241 WDYRRPPRLANFFVFLVEMGFTMFARLILISGFCDLPASASQAGITGVSHARLIFNFC 300  
 QY 301 LFEMESHSVTQAGVQWRNLGSLQPLPPGPKLFSPCSLLSSWDYGHLPHPANFCIFIRGG 360  
 Db 301 LFEMESHSVTQAGVQWRNLGSLQPLPPGPKLFSPCSLLSSWDYGHLPHPANFCIFIRGG 360  
 QY 361 VSPYLSGWSQTPDLR 375  
 Db 361 VSPYLSGWSQTPDLR 375

RESULT 10  
 ADR14409  
 ID ADR14409 standard; protein; 375 AA.  
 XX  
 AC ADR14409;  
 XX  
 DT 21-OCT-2004 (first entry)  
 XX  
 DE Human NF-kappaB pathway-associated protein SeqID410.  
 KW NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;  
 KW antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;  
 KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;  
 KW immunosuppressive; vulnary; gene therapy; immune disorder;  
 KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;  
 KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;  
 KW hyper-IgM syndrome; hypochidrotic ectodermal dysplasia;  
 KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;  
 KW viral infection; HIV-1; HIV-1; hepatitis B; hepatitis C; EBV; influenza;  
 KW viral replication; host cell survival; evasion of immune response;  
 KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;  
 KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;  
 KW autoimmune disorder; hyper immune activity;  
 KW aberrant acute phase response; hypercongenital condition; birth defect;  
 KW necrotic lesion; wound; organ transplant rejection;  
 KW aberrant signal transduction; proliferating disorder; cancer;  
 KW HIV propagation; human.  
 OS Homo sapiens.  
 PN WO2004065577-A2.  
 XX  
 PD 05-AUG-2004.  
 XX  
 PF 13-JAN-2004; 2004WO-US0000798.  
 XX  
 PR 14-JAN-2003; 2003US-0440068P.  
 PR 12-MAY-2003; 2003US-046957P.  
 XX  
 PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX  
 PI Nadler SG, Neubauer MG, Feder JN, Carman J;  
 XX  
 DR WPI; 2004-562168/54.  
 XX N-PSDB; ADR14408.  
 XX  
 PT New isolated polynucleotides and polypeptides associated with NF-kappaB  
 PT pathway, useful for diagnosing, treating, or preventing disorders or  
 XX diseases associated with NF-kappaB pathway.  
 PS Claim 6; SEQ ID NO 410; 237pp; English.  
 XX  
 CC This invention relates to the novel association of protein sequences (and  
 CC the genes which encode them) to the NF-kappaB pathway. The invention may  
 CC be useful for the production of compounds with an antiinflammatory,  
 CC cytostatic, hepatotropic, virucide, antarthritic, antirheumatic,  
 CC gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic,

CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or  
 CC vulnary activity or for gene therapy. The proteins and nucleotides are  
 CC useful for diagnosing, preventing, treating, or ameliorating conditions  
 CC or diseases associated with the NF-kappaB pathway. The condition is an  
 CC immune disorder, an inflammatory disorder, an inflammatory disorder  
 CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,  
 CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM  
 CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic  
 CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,  
 CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell  
 CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory  
 CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick  
 CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper  
 CC immune activity, disorders related to aberrant acute phase responses,  
 CC hypergenital conditions, birth defects, necrotic lesions, wounds,  
 CC organ transplant rejection, conditions related to organ transplant  
 CC rejection, disorders related to aberrant signal transduction,  
 CC proliferating disorders, cancers and HIV propagation in cells infected  
 CC with other viruses. The present sequence is that of a human protein which  
 CC is subject to the novel association with the NF-kappaB pathway of the  
 CC invention. Note: This sequence does not appear in the specification but  
 CC was obtained by the indexer from Genbank.  
 XX  
 SQ Sequence 375 AA;

Query Match 100.0%; Score 2034; DB 8; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 1e-198;  
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MEFSLLPLRLCNGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60  
 DB 1 MEFSLLPLRLCNGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60  
 QY 61 LHVGQAGLELPTSDPSVSASQSAARYTGHARLCLANFCGRNRVSLMCPSPKQST 120  
 DB 61 LHVGQAGLELPTSDPSVSASQSAARYTGHARLCLANFCGRNRVSLMCPSPKQST 120  
 QY 121 CUSLPKCDYRAAAYVGLFELFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180  
 DB 121 CUSLPKCDYRAAAYVGLFELFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180  
 QY 181 VAGTKDMHYTWLIFIFNLRQSLNSVTQAGVQWNLGSLQPLPPGKLFSCPSLSS 240  
 DB 181 VAGTKDMHYTWLIFIFNLRQSLNSVTQAGVQWNLGSLQPLPPGKLFSCPSLSS 240  
 QY 241 WDYRPPRLANFVFLVEMGFTWFAFLILISGCDLPASASQAGITGVSHHARLIFNFC 300  
 DB 241 WDYRPPRLANFVFLVEMGFTWFAFLILISGCDLPASASQAGITGVSHHARLIFNFC 300  
 QY 301 LFEMESHVSTQAGVQWNLGSLQPLPPGKLFSCPSLSS 360  
 DB 301 LFEMESHVSTQAGVQWNLGSLQPLPPGKLFSCPSLSS 360  
 QY 361 VSPYLSGWSQTPDLR 375  
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 11  
 ID AAR95913  
 AC AAR95913; protein; 397 AA.  
 XX AAR95913;  
 XX 13-NOV-1996 (first entry)  
 DT Neural thread protein.  
 DE  
 DE Neural thread protein.  
 KW Neural thread protein; NTP; diagnosis; detection; Alzheimer's disease;  
 KW neuroectodermal tumour; malignant astrocytoma; monoclonal antibody;  
 XX binding fragment.  
 XX Homo sapiens.  
 OS

XX WO9615272-A1.  
 PN 23-MAY-1996.  
 XX 14-NOV-1995; 95WO-US017111.  
 XX 14-NOV-1994; 94US-00340426.  
 XX (GEO) GEN HOSPITAL CORP.  
 XX De La Monte S, Wands JR;  
 WPI: 1996-259865/26.  
 N-PSDB; AAT27738.  
 XX Detection of neural thread protein in diagnosis of Alzheimer's disease -  
 also NTP DNA and protein sequences used in gene and anti:sense therapy.  
 PS Claim 22; Page 171-172; 238pp; English.  
 XX A method for detecting the presence of neural thread protein (NTP) having  
 a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human subject  
 comprises (a) contacting a sample from a human subject that is suspected  
 of containing the NTP with at least one molecule capable of binding to  
 the protein; and (b) detecting any of the molecule bound to the protein.  
 The binding molecule is selected from an antibody free of natural  
 impurities, a monoclonal antibody or a binding fragment of either of  
 these. The method may be used for diagnosing the presence of Alzheimer's  
 disease, neuroectodermal tumours and a malignant astrocytoma in a human  
 XX  
 SQ Sequence 397 AA;

Query Match 69.6%; Score 1415.5; DB 2; Length 397;  
 Best Local Similarity 74.6%; Pred. No. 1.7e-135;  
 Matches 285; Conservative 10; Mismatches 44; Indels 43; Gaps 7;  
 QY 1 MEFSLLPLRLCNGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60  
 DB 1 MEFSLLPLRLCNGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60  
 QY 61 LHVGQAGLELPTSDPSVSASQSAARYTGHARLCLANFCGRNRVSLMCP 110  
 DB 61 LHVGQAGLELPTSDPSVSASQSAARYTGHARLCLANFCGRNRVSLMCP 110  
 QY 111 SWSPELKQSTCLSLPKCDYRAAAYVGLFELFRLHRCPTLTQDEVQWCDHSSLOPSTPE 170  
 DB 111 SWSPELKQSTCLSLPKCDYRAAAYVGLFELFRLHRCPTLTQDEVQWCDHSSLOPSTPE 170  
 QY 171 ----IKHPPASQVAGTKDMHYTWLIFIFNLRQSLNSVTQAGVQWNLGSLQPL 226  
 DB 171 SSILPQPP----KVAGTKDMHYTWLIFIFNLRQSLNSVTQAGVQWNLGSLQPL 226  
 QY 227 PGKLFSCPSLSSWDYRPPRLANFVFLVEMGFTWFAFLILISGCDLPASASQAGITGVSHHARLIFNFC 278  
 DB 227 PGKLFSCPSLSSWDYRPPRLANFVFLVEMGFTWFAFLILISGCDLPASASQAGITGVSHHARLIFNFC 281  
 QY 279 SASQAGITGVSHHARLIFNFCLEFEMESHVSTQAGVQWNLGSLQPLPPGKLFSCPSLSS 338  
 DB 282 --PKVLGLQDVTPTARPIFNFCLEFEMESHVSTQAGVQWNLGSLQPLPPGKLFSCPSLSS 339  
 QY 339 SSWDYGHLPHPANFCIFIRGG 360  
 DB 340 SSWDYGHLPHPANFCIFIRGG 361

RESULT 12  
 ID AAU30235  
 AC AAU30235; standard; protein; 381 AA.  
 XX AAU30235;  
 XX 18-DEC-2001 (first entry)  
 DT

XX Novel human secreted protein #726.  
DE Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX Homo sapiens.  
XX WO200179449-A2.  
XX 25-OCT-2001.  
XX 16-APR-2001; 2001WO-US008656.  
XX 18-APR-2000; 2000US-00552929.  
XX 26-JAN-2001; 2001US-00770160.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT;  
XX WPI; 2001-611725/70.  
XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
XX vaccination, testing and therapy.  
XX Claim 20; Page 264-265; 765pp; English.  
XX The invention relates to novel human secreted polypeptides. The  
XX polypeptides and antibodies to the polypeptides are a disease associated  
XX determining the presence of or predisposition to a disease associated  
XX with altered levels of polypeptide. The polypeptides are also useful for  
XX identifying agents (agonists and antagonists) that bind to them. Cells  
XX expressing the proteins are useful for identifying a therapeutic agent  
XX for use in treatment of a pathology related to aberrant expression or  
XX physiological interactions of the polypeptide. Vectors comprising the  
XX nucleic acids encoding the polypeptides and cells genetically engineered  
XX to express them are also useful for producing the proteins. The proteins  
XX are useful in genetic vaccination, testing and therapy, and can be used  
XX as nutritional supplements. They may be used to increase stem cell  
XX proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
XX and/or nerve tissue growth or regeneration; immune suppression and/or  
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
XX AAU29510-AAU33304 represent the amino acid sequences of novel human  
XX secreted proteins of the invention  
XX Sequence 381 AA;  
Query Match 45.6%; Score 928; DB 4; Length 381;  
Best Local Similarity 55.9%; Pred. No. 1e-85;  
Matches 223; Conservative 27; Mismatches 101; Indels 48; Gaps 14;  
QY 4 SLLPLRLCNGAISAHNRNLRPGSSDPSASAPVAGITGMCTHARLILYFFVLEHFLHV 63  
DB 4 SLLPLRLCNGAILAHNCNLCLSGSSDPSASASQVTGKCHHTQLI-FVFLVEMGFHHI 62  
QY 64 GAGLEPLTSDPSVSASQARVTRG--HARCLANFCGRNRVSLMCPSW----- 112  
DB 63 AQAGLELLTSDPTL-ASQSAGI-TGVNHHAWLFF--FCSRDVSLCYPGWSRVAXSRIT 118  
QY 113 --SPELKQSTCLSLPKCWDRRAAV-PGLFILFPLRHRCTLTQDEVQWCDHSSLPQ--- 166  
DB 119 ATSPGLKXKXACSLPSSRDYRHVPHPGNFCIF-----GRDEVSPC-----WPGWF 164  
QY 167 STPEIKHPASASQVAGTKDMHHYTW----LIPFIFNRLQSLNSVTQAGVQWNLGSL 222  
DB 165 XTPDLRYPPASASQSAEIIIGVSHHTWPQBVFLFLNLFIYLRWLSDSVAQARVQRDLGSL 224  
QY 223 QPLPGFKLFCSESLSSWDYR-PPRLANFVFLVEMGTFWFARLILISGCDLPASAS 281  
DB 225 QAPPFRKFPKFCSLSPSSWDYRRPPPPHPANFFVFLVETGTVLARRVLISXPRDLPASAS 284

QY 282 QSAGITGVSHHARLIFNFCLEFEMESHVSQTQAGVQWPNLGLSQPLPGLKRFSCLSLPSW 341  
DB 285 QSAGITGVSHHARLIFNFP--FETGTHSVTAAVQWVTIGSLQRTPELXKSSHLILTSNW 342  
QY 342 DYGHLPHPAN-----FCIFIRGGVSPYLSGWSQTPDLR 375  
DB 343 DYRCTPPCPNPLFIYLFYFHRDEGSLCCPGWSXTPELX 381  
RESULT 13  
AAU32610  
ID AAU32610 standard; protein; 382 AA.  
XX  
AC AAU32610;  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human secreted protein #3101.  
XX  
KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX  
OS Homo sapiens.  
XX  
XX WO200179449-A2.  
XX 25-OCT-2001.  
XX 16-APR-2001; 2001WO-US008656.  
XX 18-APR-2000; 2000US-00552929.  
XX 26-JAN-2001; 2001US-00770160.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT;  
XX WPI; 2001-611725/70.  
XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
XX vaccination, testing and therapy.  
XX Claim 20; Page 637; 765pp; English.  
XX The invention relates to novel human secreted polypeptides. The  
XX polypeptides and antibodies to the polypeptides are useful for  
XX determining the presence of or predisposition to a disease associated  
XX with altered levels of polypeptide. The polypeptides are also useful for  
XX identifying agents (agonists and antagonists) that bind to them. Cells  
XX expressing the proteins are useful for identifying a therapeutic agent  
XX for use in treatment of a pathology related to aberrant expression or  
XX physiological interactions of the polypeptide. Vectors comprising the  
XX nucleic acids encoding the polypeptides and cells genetically engineered  
XX to express them are also useful for producing the proteins. The proteins  
XX are useful in genetic vaccination, testing and therapy, and can be used  
XX as nutritional supplements. They may be used to increase stem cell  
XX proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
XX and/or nerve tissue growth or regeneration; immune suppression and/or  
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
XX AAU29510-AAU33304 represent the amino acid sequences of novel human  
XX secreted proteins of the invention  
XX Sequence 382 AA;  
Query Match 45.6%; Score 927.5; DB 4; Length 382;  
Best Local Similarity 55.8%; Pred. No. 1.1e-85;  
Matches 223; Conservative 27; Mismatches 101; Indels 49; Gaps 14;  
QY 4 SLLPLRLCNGAISAHNRNLRPGSSDPSASAPVAGITGMCTHARLILYFFVLEHFLHV 63  
DB 4 SLLPLRLCNGAILAHNCNLCLSGSSDPSASASQVTGKCHHTQLI-FVFLVEMGFHHI 62



QY 64 GOAGLELPTSDPSVSASQARYRTG--HHARLCLANFCGRNRVSLMCPWS----- 113  
 Db 63 AQAGLELLTSDPTL-ASQSAGI-TGVNHHAWLFF--FCSRDVTSLCYPGWGRVAXSRIT 118  
 QY 114 ----PELKQSTCLSLPKCWDYRAAV-PGLFILFLRHRCPRTLTODEVQWCHDSSLQP-- 166  
 Db 119 ATSAPGLKXKXACFSLPSSRDYRHVPHPGFCIF-----GRDEVSPC-----WPGW 164  
 QY 167 -STPEIKHPASASQVAGTKDMHHYTW-----LIPFIFNFIQLRQSLNSVTQAGVQWRNLGS 221  
 Db 165 FXTDLRYPPASASQSAEIIQVSHHTWQVEFLFNLFIYLRWSLDSVAQARVQRDLGS 224  
 QY 222 LQPLPPGPKLFSCPSLLSSWDYRR--PRLANFFVFLVEMGFTMFARLILISGCPDLASA 280  
 Db 225 LQAPPPRFKPFSCLSLPSWDYRRPPHPANFFVLVETGTVLARRVLISXPRDLASA 284  
 QY 281 QSAGITGVSHHARLIFNFCLEFEMESHVSVTQAGVQWPNLGSLOPLPPGLKRFCSLSLPS 340  
 Db 285 QSAGITGVSHHARLIFNF--FETGTHSVTAAVQWYTTIGSLQRTPELXSSHLILTSN 342  
 QY 341 WDYGHLPPHPAN-----FCIFIRGVSPYLSGWSOTPDLR 375  
 Db 343 WDYRCTPPCPNLFYLFYFHRDEGSLCCPGWSXTPELK 382

RESULT 14  
 AAU31818  
 ID AAU31818 standard; protein; 382 AA.  
 AC AAU31818;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Novel human secreted protein #2309.  
 XX  
 KW Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200179449-A2.  
 XX  
 PD 25-OCT-2001.  
 XX  
 PF 16-APR-2001; 2001WO-US008656.  
 XX  
 PR 18-APR-2000; 2000US-00552929.  
 PR 26-JAN-2001; 2001US-00770160.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Liu C, Drmanac RT;  
 XX WPI; 2001-611725/70.  
 XX  
 PT Nucleic acids encoding a range of human polypeptides, useful in genetic  
 PT vaccination, testing and therapy.  
 XX  
 PS Claim 20; Page 513; 765pp; English.  
 XX  
 CC The invention relates to novel human secreted polypeptides. The  
 CC polypeptides and antibodies to the polypeptides are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of polypeptide. The polypeptides are also useful for  
 CC identifying agents (agonists and antagonists) that bind to them. Cells  
 CC expressing the proteins are useful for identifying a therapeutic agent  
 CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell

CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention  
 XX  
 SQ Sequence 382 AA;  
 Query Match 45.6%; Score 927.5; DB 4; Length 382;  
 Best Local Similarity 55.8%; Pred. No. 1.1e-85;  
 Matches 223; Conservative 27; Mismatches 101; Indels 49; Gaps 14;  
 QY 4 SLLPRLCNGAISHRNLRLLPGSSDSPASAPVAGITGMCTHARLILLYFLVEMFLHV 63  
 Db 4 SLLPRLCNGAILAHCLNCLSGSSDSPASASQVGTGKCHHTQLI-FVFLVEMGFHHI 62  
 QY 64 GOAGLELPTSDPSVSASQARYRTG--HHARLCLANFCGRNRVSLMCPWS----- 113  
 Db 63 AQAGLELLTSDPTL-ASQSAGI-TGVNHHAWLFF--FCSRDVTSLCYPGWGRVAXSRIT 118  
 QY 114 ----PELKQSTCLSLPKCWDYRAAV-PGLFILFLRHRCPRTLTODEVQWCHDSSLQP-- 166  
 Db 119 ATSAPGLKXKXACFSLPSSRDYRHVPHPGFCIF-----GRDEVSPC-----WPGW 164  
 QY 167 -STPEIKHPASASQVAGTKDMHHYTW-----LIPFIFNFIQLRQSLNSVTQAGVQWRNLGS 221  
 Db 165 FXTDLRYPPASASQSAEIIQVSHHTWQVEFLFNLFIYLRWSLDSVAQARVQRDLGS 224  
 QY 222 LQPLPPGPKLFSCPSLLSSWDYRR--PRLANFFVFLVEMGFTMFARLILISGCPDLASA 280  
 Db 225 LQAPPPRFKPFSCLSLPSWDYRRPPHPANFFVLVETGTVLARRVLISXPRDLASA 284  
 QY 281 QSAGITGVSHHARLIFNFCLEFEMESHVSVTQAGVQWPNLGSLOPLPPGLKRFCSLSLPS 340  
 Db 285 QSAGITGVSHHARLIFNF--FETGTHSVTAAVQWYTTIGSLQRTPELXSSHLILTSN 342  
 QY 341 WDYGHLPPHPAN-----FCIFIRGVSPYLSGWSOTPDLR 375  
 Db 343 WDYRCTPPCPNLFYLFYFHRDEGSLCCPGWSXTPELK 382  
 RESULT 15  
 AAU32707  
 ID AAU32707 standard; protein; 382 AA.  
 XX  
 AC AAU32707;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Novel human secreted protein #3198.  
 XX  
 KW Human; vaccination; gene therapy; nutritional supplement;  
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200179449-A2.  
 XX  
 PD 25-OCT-2001.  
 XX  
 PF 16-APR-2001; 2001WO-US008656.  
 XX  
 PR 18-APR-2000; 2000US-00552929.  
 PR 26-JAN-2001; 2001US-00770160.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Liu C, Drmanac RT;  
 XX WPI; 2001-611725/70.  
 XX  
 PT Nucleic acids encoding a range of human polypeptides, useful in genetic

PT vaccination, testing and therapy.  
XX  
PS Claim 20; Page 652; 765pp; English.  
XX  
XX The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used  
CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
CC secreted proteins of the invention  
XX  
SQ Sequence 382 AA;

Query Match 45.6%; Score 927.5; DB 4; Length 382;  
Best Local Similarity 55.8%; Pred. No. 1.1e-85;  
Matches 223; Conservative 27; Mismatches 101; Indels 49; Gaps 14;  
QY 4 SLLPLECNGALSAHNRNLPGSSDSPASAPVAGTCMCTHARLILYFFLVEHFLHV 63  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :  
4 SLLPLECNGALSAHNRNLPGSSDSPASAPVAGTCMCTHARLILYFFLVEHFLHV 62  
QY 64 GAGLELPTSDPSVSASQARVGTG--HARLCLANFCGRNRVSLMCPSPWS----- 113  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :  
63 AQAGLELLTSDFTL-ASQAGI-TGVNHWLFF--FCSDTVSLCYFGNSRVAXSRIT 118  
QY 114 ----PELKOSTCLSLPKWDYRAAV-PGLFILFLRHRCPRTLTODEVQWCDHSSLP-- 166  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :  
119 ATSAPGLKXKXCFSLPSSRDYRHPHPGNCIF-----GRDEVSPC-----WPGW 164  
QY 167 -STPEIKHPASQVAGTKDMHHTW----LIFIFNPLRQSLNSVTOAGVQWRNLGS 221  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :  
165 FXTDLYPPASASQSAEIIQVSHHTWQVEFLFLMLFYLRWSLDSVAQARVQRDLGS 224  
QY 222 LQPLPFGKLFSCPSLSSWDYR--PPRLANFFVFLVEMGFTWEARLILISGCDLPASA 280  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :  
225 LQAPPRFPKFCCLPSSWDYRPPHPANFFVFLVETGFTVLARRVLISAPRDLPSA 284  
QY 281 SQSAGITGVSHARLIFNFCLEFEMESHVTOAGVQWPNIGSLQPLPPGLKRFSCLSLPS 340  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :  
285 SQSAGITGVSHHTRLIFNF--FETGTHSVTAAVQWYTIQSLQPRTPELKXSSHLITSN 342  
QY 341 WDYGHLPHPAN-----FCIFIRGGVSPYLSGWSQTPDLR 375  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :  
343 WDYRCTPFCPPNLIYFLFYFHRDEGSLCCPGNSXTPELK 382



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 15, 2005, 22:16:29 ; Search time 810 Seconds  
(without alignments)  
3084.524 Million cell updates/sec

Title: US-09-380-203-2  
Perfect score: 2034  
Sequence: 1 MEFSLLPLRLCNGAISAHR.....FIRGGVSPYLSGWSQTFDLR 375

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 7389322 seqs, 331285599 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Post-processing: Minimum Match 0%

Command line parameters: -MODEL=frame\_p2n.model -DRV=xlh

-Q=/cgn2\_1/USPTO.spool/US09380203/runat\_13092005\_170811\_11699/app\_query.fasta\_1.519  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cgi -LIST=45 -LOCAL -OUTFMT=ptc -NORM=ext -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09380203@cgn\_1\_1\_800@runat\_13092005\_170811\_11699  
-NCPU=6 -FCPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2034	100.0	1442	9	US-09-964-666-1 Sequence 1, Appli
2	2034	100.0	1442	9	US-09-964-412-1 Sequence 1, Appli
3	2034	100.0	1442	10	US-09-964-667-1 Sequence 1, Appli
4	2034	100.0	1442	10	US-09-872-968-1 Sequence 1, Appli
5	2034	100.0	1442	10	US-09-964-678A-1 Sequence 1, Appli
6	2034	100.0	1442	14	US-10-146-130-1 Sequence 1, Appli
7	2034	100.0	1442	14	US-10-092-934-1 Sequence 1, Appli
8	2034	100.0	1442	14	US-10-153-334-53 Sequence 53, Appli
9	2034	100.0	1442	14	US-10-198-069-48 Sequence 48, Appli
10	2034	100.0	1442	15	US-10-198-070-125 Sequence 125, App
11	2034	100.0	1442	19	US-10-755-889-409 Sequence 409, App
12	2034	100.0	1442	21	US-10-910-173-1 Sequence 1, Appli
13	1787	87.9	1418	9	US-09-964-666-4 Sequence 4, Appli
14	1787	87.9	1418	9	US-09-964-412-4 Sequence 4, Appli
15	1787	87.9	1418	10	US-09-964-667-4 Sequence 4, Appli
16	1787	87.9	1418	10	US-09-964-678A-4 Sequence 4, Appli
17	1520	74.7	1381	9	US-09-964-666-3 Sequence 3, Appli
18	1520	74.7	1381	9	US-09-964-412-3 Sequence 3, Appli
19	1520	74.7	1381	10	US-09-964-667-3 Sequence 3, Appli
20	1520	74.7	1381	10	US-09-964-678A-3 Sequence 3, Appli
21	964.5	47.4	75252	13	US-10-087-192-904 Sequence 904, App
C 22	901	44.3	33362	19	US-10-388-838-36 Sequence 36, Appli
C 23	892.5	43.9	92638	19	US-10-450-826-3 Sequence 3, Appli
C 24	871.5	42.8	103574	22	US-10-756-149-2307 Sequence 2307, Ap
25	870	42.8	193691	20	US-10-719-993-6768 Sequence 6768, Ap
C 26	870	42.8	18760	20	US-10-719-993-6765 Sequence 1542, Ap
27	863.5	42.5	110096	9	US-09-880-107-1542 Sequence 208, App
C 28	861.5	42.4	21666	18	US-10-052-482-208 Sequence 13, Appli
C 29	858	42.2	126001	17	US-10-175-492-13 Sequence 3, Appli
30	857	42.1	32865	13	US-10-087-192-964 Sequence 964, App
31	851.5	41.9	3379	16	US-10-268-332-3 Sequence 3, Appli
32	851.5	41.9	3379	20	US-10-775-965-3 Sequence 3, Appli
33	851.5	41.9	28953	20	US-10-719-993-6811 Sequence 6811, Ap
34	851.5	41.7	15783	13	US-09-764-872-700 Sequence 700, App
35	848.5	41.7	61103	13	US-10-087-192-58 Sequence 58, Appli
36	848	41.6	124990	20	US-10-684-422-156 Sequence 156, App
C 37	846	41.4	40645	9	US-09-818-656A-3 Sequence 3, Appli
38	841.5	41.4	40645	13	US-10-216-441-3 Sequence 3, Appli
39	841.5	41.4	40645	22	US-10-961-230-3 Sequence 3, Appli
40	841.5	41.4	40645	9	US-09-764-860-1046 Sequence 1046, Ap
41	840	41.3	27148	14	US-10-074-095-1046 Sequence 1046, Ap
42	840	41.3	27148	17	US-10-212-872-1046 Sequence 1120, Ap
43	840	41.3	74234	13	US-10-087-192-1120 Sequence 316, App
44	840	41.3	74234	13	US-10-087-192-1120 Sequence 316, App
45	839.5	41.3	45855	17	US-10-085-117-316 Sequence 316, App

ALIGNMENTS

RESULT 1  
US-09-964-666-1  
; Sequence 1, Application US/09964666  
; Patent No. US20020104108A1  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; Wands, Jack R.  
; TITLE OF INVENTION: Transgenic Animals and Cell Lines for  
; Screening Drugs Effective for the Treatment or Prevention  
; of Alzheimer's Disease  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.  
; STREET: 1100 New York Ave., Suite 600  
; CITY: Washington  
; STATE: DC



```

675 Db TCACCTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCTCGCCGCCAGCCTCCTGAGTAGC 734
241 QY TTPAspTyrArgArgProProArgLeuAlaAasnPhePheValPheLeuValGluMetGly 260
735 Db TGGAGCTACAGGCGCCACACACGCTAGCTAAATTTTTTTGTATTTTTTTAGTAGAGATGGG 794
261 QY PheThrMetPheAlaArgLeuLeuLeuLeuSerGlyProCysAspLeuProAlaSerAla 280
795 Db TTCACCATGTTGCCAGGTGATCTTGATCTCTGGACCTTGTCATCTGCCTGCGCTCGGCC 854
281 QY SerGlnSerAlaGlyIleThrGlyValSerHisAlaArgLeuLeuLeuPheAasnPheCys 300
855 Db TCCCAAGTGTGGGATTACAGCGGTGAGCCACACCGCGCTTAATTTTTTAAATTTTGT 914
301 QY LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnThrProAasnLeuGly 320
915 Db TTGTTTGAATGAATGAACTCACTCTGTACCCAGCTGGAGTGCATGGCCAAATCTCGGC 974
321 QY SerLeuGlnProLeuProProGlyLeuLeuLysArgPheSerCysLeuSerLeuProSerSer 340
975 Db TCACCTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCTGTCTCAGCCTCCCAAGCAGC 1034
341 QY TTPAspTyrGlyHisLeuProProHisProAlaAasnPheCysIlePheIleArgGlyGly 360
1035 Db TGGGATTACGGGCACCTGCACCACACCCCGCTAAATTTTTTGTATTTTTCATTAGAGCGGG 1094
361 QY ValSerProTyrLeuSerGlyTyrSerGlnThrProAasnLeuArg 375
1095 Db GTTTCACCAATATTGCAGCGCTGCTCAAACTCTCGACCTCAGG 1139

RESULT 3
US-09-964-667-1
: Sequence 1, Application US/09964667
: Publication No. US20030033621A1
: GENERAL INFORMATION:
: APPLICANT: de la Monte, Suzanne
: Wanda, Jack R.
: TITLE OF INVENTION: Transgenic Animals and Cell Lines for
: Screening Drugs Effective for the Treatment or Pre-
: of Alzheimer's Disease
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
: STREET: 1100 New York Ave., Suite 600
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20005-3934
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION NUMBER: US/09/964,667
: APPLICATION NUMBER: 32,893
: FILING DATE: 28-Sep-2001
: CLASSIFICATION: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Esmond, Robert W.
: REGISTRATION NUMBER: 32,893
: REFERENCE/DOCKET NUMBER: 0609.4370000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-371-2600
: TELEFAX: 202-371-2540
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1442 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: both
: MOLECULE TYPE: cDNA
: FEATURE:

```

```
; NAME/KEY: CDS
; LOCATION: 15..1139
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-964-667-1

Alignment Scores:
Pred. No.: 7,83e-197 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-380-203-2 (1-375) x US-09-964-667-1 (1-1442)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
DB 15 ATGGAGTTTTCGCTCTTGTGGCCAGGCTGGAGTGCAATGGCGCAATCTCAGCTCACC GC 74

QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
DB 75 AACCTCCGCTCCCGGGTTCACAGCGATTCTCTGCGCTCAGCCTCCCGAGTAGCTGGGATT 134

QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60
DB 135 ACAGGATGTGCACCCACGCTCGGCTAAATTTTGTATTTTGTATTTTGTATTTTGTATTT 194

QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspSerProSerValSerAla 80
DB 195 CTCCTATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATCCCTCCGCTCGG GC 254

QY 81 SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys 100
DB 255 TCCCAAAAGTGTAGATACAGGACTGGCCACCATGGCCGCTCTGCTGGCTAAATTTTGT 314

QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120
DB 315 GGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGCTCAAGCAGTCCACC 374

QY 121 CysLeuSerLeuProLysCysTrpAspTyrArgAlaAlaValProGlyLeuPheIle 140
DB 375 TGCCTCAGCCTCCCAAGTGTGGGATTACAGCGGTGACGCGTGCAGCTGCTGCTTTTATT 434

QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
DB 435 TTATTTTATTAAGACACAGGTGTCCACTCTTACCACAGATGAAGTGCAGTGGTGTGAT 494

QY 161 HisSerSerLeuGlnProSerThrProGluIleLysHisProAlaSerAlaSerGln 180
DB 495 CACAGCTCACTCAGCGCTTCACTCTGAGATCAGCATCTCTCTGCTCAGCTCAGCTCCCAA 554

QY 181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePhePheIlePheAsn 200
DB 555 GTAGCTGGGACCAAAAGACATGCACCACTACACTGCTGGCTAAATTTTATTTTATTTTAA 614

QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
DB 615 TTTTGTAGACAGAGTCTCAACTCTGFCACCCAGGCTGGAGTGCAGTGGCGCAATCTTGGC 674

QY 221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
DB 675 TCACATGCACACTCTGCTCCCGGTTCAAGTTATTTCTCTGCTCCCGACGCTCTCTGAGTAGC 734

QY 241 TrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeuValGluMetGly 260
DB 735 TGGGACTACAGGCGGCCACAGCGCTAGCTAAATTTTGTATTTTGTATTTTGTAGAGATGGG 794

QY 261 PheThrMetPheAlaArgLeuIleLeuSerGlyProCysAspLeuProAlaSerAla 280
DB 795 TTCACCATGTTGCCAGGTTGATCTTGATCTCTGGACCTTGATCTGCTGCTGCTGCTGCTG 854

QY 281 SerGlnSerAlaGlyIleThrGlyValSerHisAlaArgLeuIlePheAsnPheCys 300
DB 854 TCCCAAAAGTGTAGATACAGGACTGGCCACCATGGCCGCTCTGCTGGCTAAATTTTGT 914

QY 301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320
DB 915 TTGTTTGAATGGAATCTCACTCTGTATCCCGAGGCTGGAGTGAATGGCAAATCTCGGC 974

QY 321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340
DB 975 TCACATGCACACTCTGCTCCCGGCTCAAGGATTTCTCTGTCTCAGCTCCCAAGCAGC 1034

QY 341 TrpAspTyrGlyHisLeuProHisProAlaAsnPheCysIlePheIleArgGlyGly 360
DB 1035 TGGGATTACGGGCACCTGCCACACACCCCGCTAAATTTTGTATTTTGTATTTTGTATTT 1094

QY 361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375
DB 1095 GTTTCACCATATTTTGTAGGCTGGTCTCAAACTCTCAGCTCAGCTCAGG 1139

RESULT 4
US-09-872-968-1
; Sequence 1, Application US/09872968
; Publication No. US20030050262A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R
; APPLICANT: de la Monte, Suzanne M
; TITLE OF INVENTION: Inhibition of Neurodegeneration
; FILE REFERENCE: 21486-047
; CURRENT APPLICATION NUMBER: US/09/872,968
; CURRENT FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-872-968-1

Alignment Scores:
Pred. No.: 7,83e-197 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-380-203-2 (1-375) x US-09-872-968-1 (1-1442)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
DB 15 ATGGAGTTTTCGCTCTTGTGGCCAGGCTGGAGTGCAATGGCGCAATCTCAGCTCACC GC 74

QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
DB 75 AACCTCCGCTCCCGGGTTCACAGCGATTCTCTGCGCTCAGCCTCCCGAGTAGCTGGGATT 134

QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60
DB 135 ACAGGATGTGCACCCACGCTCGGCTAAATTTTGTATTTTGTATTTTGTATTTTGTATTT 194

QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspSerProSerValSerAla 80
DB 195 CTCCTATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATCCCTCCGCTCGG GC 254

QY 81 SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys 100
DB 255 TCCCAAAAGTGTAGATACAGGACTGGCCACCATGGCCGCTCTGCTGGCTAAATTTTGT 314

QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120
DB 315 GGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGCTCAAGCAGTCCACC 374

QY 121 CysLeuSerLeuProLysCysTrpAspTyrArgAlaAlaValProGlyLeuPheIle 140
DB 375 TGCCTCAGCCTCCCAAGTGTGGGATTACAGCGGTGACGCGTGCAGCTGCTGCTTTTATT 434

QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
DB 435 TTATTTTATTAAGACACAGGTGTCCACTCTTACCACAGATGAAGTGCAGTGGTGTGAT 494

QY 161 HisSerSerLeuGlnProSerThrProGluIleLysHisProAlaSerAlaSerGln 180
DB 495 CACAGCTCACTCAGCGCTTCACTCTGAGATCAGCATCTCTCTGCTCAGCTCAGCTCCCAA 554

QY 181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePhePheIlePheAsn 200
DB 555 GTAGCTGGGACCAAAAGACATGCACCACTACACTGCTGGCTAAATTTTATTTTATTTTAA 614

QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
DB 615 TTTTGTAGACAGAGTCTCAACTCTGFCACCCAGGCTGGAGTGCAGTGGCGCAATCTTGGC 674

QY 221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
DB 675 TCACATGCACACTCTGCTCCCGGTTCAAGTTATTTCTCTGCTCCCGACGCTCTCTGAGTAGC 734

QY 241 TrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeuValGluMetGly 260
DB 735 TGGGACTACAGGCGGCCACAGCGCTAGCTAAATTTTGTATTTTGTATTTTGTAGAGATGGG 794

QY 261 PheThrMetPheAlaArgLeuIleLeuSerGlyProCysAspLeuProAlaSerAla 280
DB 795 TTCACCATGTTGCCAGGTTGATCTTGATCTCTGGACCTTGATCTGCTGCTGCTGCTGCTG 854

QY 281 SerGlnSerAlaGlyIleThrGlyValSerHisAlaArgLeuIlePheAsnPheCys 300
DB 854 TCCCAAAAGTGTAGATACAGGACTGGCCACCATGGCCGCTCTGCTGGCTAAATTTTGT 914

QY 301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320
DB 915 TTGTTTGAATGGAATCTCACTCTGTATCCCGAGGCTGGAGTGAATGGCAAATCTCGGC 974

QY 321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340
DB 975 TCACATGCACACTCTGCTCCCGGCTCAAGGATTTCTCTGTCTCAGCTCCCAAGCAGC 1034

QY 341 TrpAspTyrGlyHisLeuProHisProAlaAsnPheCysIlePheIleArgGlyGly 360
DB 1035 TGGGATTACGGGCACCTGCCACACACCCCGCTAAATTTTGTATTTTGTATTTTGTATTT 1094

QY 361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375
DB 1095 GTTTCACCATATTTTGTAGGCTGGTCTCAAACTCTCAGCTCAGCTCAGG 1139
```



```

375  TGCTCAGCGCTCCCAAGTGTGGGATTACAGCGGTGACGGTGTGCTGGCGCTTTTATT 434
QY  141  LeuPheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
Db  435  TTTATTTTATTTTAAACACAGGTGTCCCACTCTTACCAGAGTGAAGTGCAGTGGTGTGAT 494
QY  161  HisSerSerLeuGlnProSerThrProGluLeuLeuHisProProAlaSerAlaSerGln 180
Db  495  CACAGCTCACTGAGCGCTTCAACTCTGAGATCAAGCATCTCTGCTCAGCGCTCCCAA 554
QY  181  ValAlaGlyThrLysAspMetHisHisThrTrpLeuLeuPheLeuPheLeuPheAsn 200
Db  555  GTAGCTGGGACCAAGACATGACCACTACCTGCTGCTAAATTTTATTTTATTTTAAAT 614
QY  201  PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db  615  TTTTGTAGACAGAGTCTCAACTCTGTACCCAGGCTGGAGTGCAGTGGCGCAATCTTGGC 674
QY  221  SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
Db  675  TCAGTCAACCTCTGCTCCCGGTTCAAGTTATTCTCTGCCCGCAGCGCTCTGAGTAGC 734
QY  241  TrpAspTyrArgArgProArgLeuAlaAsnPheValPheLeuValGluMetGly 260
Db  735  TGGGACTACAGCGCGCCACCGCTAGCTAATTTTGTATTTTGTATTTTGTAGAGATGGG 794
QY  261  PheThrMetPheAlaArgLeuLeuLeuSerGlyProCysAspLeuProAlaSerAla 280
Db  795  TTCACCATGTTGCCAGGTGATCTTGATCTCTGGACCTTGATCTGCTCGCTCGCGCC 854
QY  281  SerGlnSerAlaGlyLeuThrGlyValSerHisHisAlaArgLeuLeuPheAsnPheCys 300
Db  855  TCCCAAGTGTGGGATTACAGCGGTGAGCCACCGCCGCTTATTTTATTTTGT 914
QY  301  LeuPheGluMetGlnSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320
Db  915  TTGTTGAATGAATCTCACTCTGTACCCAGGCTGGAGTGAATGGCAATCTCGGC 974
QY  321  SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340
Db  975  TCAGTCAACCTCTGCTCCCGGCTCAAGCGATCTCTGCTCAGCGCTCCCAAGCAGC 1034
QY  341  TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheLeuArgGlyGly 360
Db  1035  TGGGATTACGGGCACTGCCACACACACCGCTTATTTGTATTTTCAATTAGAGCGGG 1094
QY  361  ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375
Db  1095  GTTTCACCATATTTGTAGGCTGGTCTCAAACTCTCTGACCTCAGG 1139

```

RESULT 5

```

US-09-964-678A-1
; Sequence 1, Application US/09964678A
; Publication NO. US2003066097A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R.
; APPLICANT: Wands, Suzanne
; TITLE OF INVENTION: Transgenic Animals and Cell Lines for Screening Drugs
; TITLE OF INVENTION: Effective for the Treatment or Prevention of
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 0609.4370002
; CURRENT APPLICATION NUMBER: US/09/964,678A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 09/380,203
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: PCT/US98/03685
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/038,908
; PRIOR FILING DATE: 1997-02-26
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1442

```

```

; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: AD7c-NTP CDNA
; NAME/KEY: CDS
; LOCATION: (15)..(1139)
; OTHER INFORMATION:
US-09-964-678A-1

Alignment Scores:
Pred. No.: 7,83e-197 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-380-203-2 (1-375) x US-09-964-678A-1 (1-1442)
QY  1  MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyValAlaIleSerAlaHisArg 20
Db  15  ATGGAGTTTTCGCTCTTGTGTTGCCAGGCTGGAGTGCATGGCGCAATCTCAGCTCACCGC 74
QY  21  AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
Db  75  AACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCTCAGCGCTCCCGAGTAGCTGGATT 134
QY  41  ThrGlyMetCysThrHisAlaArgLeuLeuLeuTyrPhePheLeuValGluMetGluPhe 60
Db  135  ACAGCATGTGCACCGCTCGGCTAATTTGTATTTTATTTATTTAGTAGAGATGGAGTTT 194
QY  61  LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAla 80
Db  195  CTCCATGTTGTCAGGCTGGTCTCGAATCTCCGACTCAGATGATCCCTCCGCTCTCGGCC 254
QY  81  SerGlnSerAlaArgTyrArgThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCys 100
Db  255  TCCCAAGTGTGTAGATACAGGACTGGCCACCATGCCGCTCTGCTGGCTAATTTTGT 314
QY  101  GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120
Db  315  GGTAGAAAACAGGGTTTCACTGATGTGCCCAAGCTGCTCTCTGAGCTCAAGCAGTCCACC 374
QY  121  CysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheIle 140
Db  375  TGCCTCAGCGCTCCCAAGTGTGGGATTACAGCGGTGACGGCTGCGCTGGCGCTTTTAT 434
QY  141  LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
Db  435  TTTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAGTGCAGTGGTGTGAT 494
QY  161  HisSerSerLeuGlnProSerThrProGluLeuLysHisProProAlaSerAlaSerGln 180
Db  495  CACAGCTCACTGAGCGCTTCAACTCTGAGATCAAGCATCTCTCTGCTGCTCAGCGCTCCCAA 554
QY  181  ValAlaGlyThrLysAspMetHisHisThrTrpLeuLeuPheLeuPheLeuPheAsn 200
Db  555  GTAGCTGGGACCAAGACATGACCACTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 614
QY  201  PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db  615  TTTTGTAGACAGAGTCTCAACTCTGTACCCAGGCTGGAGTGCAGTGGCGCAATCTTGGC 674
QY  221  SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
Db  675  TCAGTCAACCTCTGCTCCCGGTTCAAGTTATTCTCTGCCCGCAGCGCTCTGAGTAGC 734
QY  241  TrpAspTyrArgArgProArgLeuAlaAsnPheValPheLeuValGluMetGly 260
Db  735  TGGGACTACAGCGCGCCACCGCTAGCTAATTTTGTATTTTGTATTTTGTAGAGATGGG 794
QY  261  PheThrMetPheAlaArgLeuLeuLeuSerGlyProCysAspLeuProAlaSerAla 280

```

Db	795	TTACACATGTCGCAGCGTTGATCTTGATCTCTCGACCTTGCTGATCTGCCTCGCCTCGGCC	854
Qy	281	SerGlnSerAlaGlyIleThrGlyValSerHisAlaArgLeuIlePheAsnPheCys	300
Db	855	TCCCAAAAGTGTGGGATTCACAGGCGTCAGCCACCACCGCCGCTATTTTAAATTTTGT	914
Qy	301	LeuPheGluMetGluSerHisSerValThrClnAlaGlyValGlnTrpProAsnLeuGly	320
Db	915	TTGTTTGAAATGGGAATCTCACTCTGTACCCAGGCTGGAGTGCAAATGCCAAATCTCGGC	974
Qy	321	SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer	340
Db	975	TCACGTCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCTGTCTCAGCCTCCCAAGCAGC	1034
Qy	341	TrpAspTyrGlyHisLeuProProHisIleProAlaAsnPheCysIlePheIleArgGlyGly	360
Db	1035	TGGGATTCAGGGGCACCTGCCACCAACCCCGCTAAATTTTGTATTTTCAATTAGAGGGCGG	1094
Qy	361	ValSerProTyrLeuSerGlyTyrPserGlnThrProAspLeuArg	375
Db	1095	GTTTCACATATTTGTCAAGCTGGTCTCAAACTCCTGACCTCAGG	1139

Qy	101	GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr	120
Db	315	GGTAGAACACAGGGTTTCACTGATGTGCCAACGTGGTCTCCTGAGCTCAACAGCAGTCCACC	374
Qy	121	CysLeuSerLeuProLysCysTrpAspTyrArgAlaAlaValProGlyLeuPheIle	140
Db	375	TGCCTCAGCCTCCCAAGTCTGGGATTACAGCGGTGCAGCGGTGCTGGCCCTTTTATT	434
Qy	141	LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp	160
Db	435	TTATTTTTTAAAGACACAGGTGTCCCACTCTTTACCCAGGATGAAGTGCAGTGGTGTGAT	494
Qy	161	HisSerSerLeuGlnProSerThrProGluLileIysHisProProAlaSerAlaSerGln	180
Db	495	CACAGCTCACTCAGCGCTTCAACTCTCTGAGATCAAGCATCCTCCTGCTCAGGCTCCCAA	554
Qy	181	ValAlaGlyThrLysAspMetHisHisTyrThrTripleLillePheIlePheAsn	200
Db	555	GTAGCTGGACCAAGACATGCACCATACACTGGCTAATTTTATTTTATTTTAT	614
Qy	201	PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly	220
Db	615	TTTTTTGAGACAGAGTCTCAACTCTGTCAACCAGCTGGAGTGCAGTGGCGCAATCTTGGC	674
Qy	221	SerLeuGlnProLeuProProGlyLysPheLysLeuPheSerCysProSerLeuLeuSerSer	240
Db	675	TCACTGCAACTCTGCCCTCCGGGTTCAGGTAAATCTCTGCCCCAGCTCTCTGAGTAGC	734
Qy	241	TrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeuValGluMetGly	260
Db	735	TGGGACTACAGGCGCCACCAACGCTAGCTAAATTTTTTTGTATTTTTTTAGTAGAGTGGG	794
Qy	261	PheThrMetPheAlaArgLeuLilleLeuLysSerGlyProCysAspLeuProAlaSerAla	280
Db	795	TTCAACATGTCGCCAGGTGTGATCTCTGGACCTTGTGATCTGCCTGCCTCGGCC	854
Qy	281	SerGlnSerAlaGlyLilleThrGlyValSerHisAlaArgLeuLillePheAsnPheCys	300
Db	855	TCCCAAAGTCTGGGATTACAGCGGTGAGCCACAGCCCGGCTTAATTTTAAATTTTGT	914
Qy	301	LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly	320
Db	915	TTGTTTTGAATGGAACTCTCACTCTGTATCCAGGCTGGAGTGCAAATGGCCAAATCTCGGC	974
Qy	321	SerLeuGlnProLeuProProGlyLysLysArgPheSerCysLeuSerLeuProSerSer	340
Db	975	TCACGTGCAACTCTGCCCTCCCGGCTCAACGGAATCTCTGTCTCAGCTCTCCCAAGCAGC	1034
Qy	341	TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGlyGly	360
Db	1035	TGGGATTACGGGCACCTGCCACCAACCCCGCTAAATTTTTTGTATTTTCAATTAGAGGGGG	1094
Qy	361	ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg	375
Db	1095	GTTTTCACATATTTGTCCAGCTGGTCTCAAACTCTCTGACCTCAGG	1139

RESULT 7  
US-10-092-934-1  
; Sequence 1, Application US/10092934  
; Publication No. US20030054990A1  
; GENERAL INFORMATION:  
; APPLICANT: AVERBACK, PAUL  
; TITLE OF INVENTION: METHODS OF USING NEURAL THREAD PROTEINS TO TREAT TUMORS  
; TITLE OF INVENTION: AND CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF  
; TITLE OF INVENTION: CELLS  
; FILE REFERENCE: 018792-0199  
; CURRENT APPLICATION NUMBER: US/10/092,934  
; CURRENT FILING DATE: 2002-06-05  
; PRIOR APPLICATION NUMBER: 60/273,957  
; PRIOR FILING DATE: 2001-03-08  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.1

```
; SEQ ID NO 1
; LENGTH: 1442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15)..(1139)
US-10-092-934-1

Alignment Scores:
Pred. No.: 7.83e-197 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-380-203-2 (1-375) x US-10-092-934-1 (1-1442)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
DB 15 ATGGAGTTTTCGCTCTTTGGCCAGGCTGGAGTGCATGGCGCAATCTCAGCTACCGC 74
QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
DB 75 AACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCTCAGCCTCCCGCAGTAGCTGGATT 134
QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrrPhePheLeuValGluMetGluPhe 60
DB 135 ACAGGATGTGCACCCAGCTCGGCTAAATTTTGTATTTTATTTAGTAGAGTGGATT 194
QY 61 LeuHisValGlyGlnAlaGlyLeuProThrSerAspSerProSerValSerAla 80
DB 195 CTCATGTGGTCAGGCTGTCTCGAATCTCCGACCTCAGATGATCCCTCGTCTCGGCC 254
QY 81 SerGlnSerAlaArgTyrrArgThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCys 100
DB 255 TCCCAAGTGTAGATACAGGACTGGCCACCATGCGCGCTCTGCTGCTAAATTTTGT 314
QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120
DB 315 GSTAGAAACAGGGTTTTCACGTATGTCCTCCCAAGCTGCTCTCCTAGCTCAAGCAGTCCACC 374
QY 121 CysLeuSerLeuProLysCysTyrAspTyrArgAlaAlaValProGlyLeuPheIle 140
DB 375 TGCTCAGCTCCCAAGTGTGGATTACAGCGTGCAGCGTGCAGCGTCTGCTGCTTTTAT 434
QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
DB 435 TTATTTTAAAGACACAGGTGTCCTCCACTCTTTACCCAGGATGAAGTGCAGTGGTGTGAT 494
QY 161 HisSerSerLeuGlnProSerThrProGluIleLysHisProProAlaSerAlaSerGln 180
DB 495 CACAGCTCACTGCAGCGCTTCACTCTCTGAGATCAAGCATCTCTGCTCAGCCTCCCAA 554
QY 181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIlePheAsn 200
DB 555 GTAGCTGGGACCAAGACATGCACCATCAGCTGGCTAAATTTTATTTTATTTTAT 614
QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
DB 615 TTTTTCAGACAGAGTCTCAACTCTGTCCACCCAGCTGGAGTGCAGTGGCGCAATCTTGGC 674
QY 221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
DB 675 TCACGTCAACTCTGCTCCCGGTTCAAGTTATTTCTCTGCCCGCAGCTCTCTAGTAGC 734
QY 241 TrpAspTyrArgArgProProArgLeuAlaAsnPheValPheLeuValGluMetGly 260
DB 735 TGGAGCTACAGGCGCCACCCAGCCTAGCTAAATTTTGTATTTTATTTAGTAGAGTGGG 794
QY 261 PheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeuProAlaSerAla 280
```

```
DB 795 TTCACCATGTTCCGCCAGGTGATCTTGATCTCTGACCTTGTGATCTGCTCGCTCGGCC 854
QY 281 SerGlnSerAlaGlyIleThrGlyValSerHisHisAlaArgLeuIlePheAsnPheCys 300
DB 855 TCCCAAAGTGTGGGATTACAGCGTGAGCCACCGCGCTTATTTTAAATTTTGT 914
QY 301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320
DB 915 TTGTTTGAATGAATCTCACTCTGTACCCAGGCTGGAGTGCATGGCCAAATCTCGGC 974
QY 321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340
DB 975 TCACGTGCAACCTCTGCTCCCGGCTCAAGCGATTCTCTGCTCAGCCTCCCAAGCAGC 1034
QY 341 TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGlyGly 360
DB 1035 TGGGATTACGGCACCTGCCACCAACCCCGCTAAATTTTGTATTTTATTTAGAGCGGG 1094
QY 361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375
DB 1095 GTTTCACCATATTTGTGAGGCTGGTCTCAAACTCTGACCTCAGG 1139

RESULT 8
US-10-153-334-53
; Sequence 53, Application US/10153334
; Publication No. US20030096350A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 59003-000006
; CURRENT APPLICATION NUMBER: US/10/153,334
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,156
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 1442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15)..(1139)
US-10-153-334-53

Alignment Scores:
Pred. No.: 7.83e-197 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-380-203-2 (1-375) x US-10-153-334-53 (1-1442)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
DB 15 ATGGAGTTTTCGCTCTTTGGCCAGGCTGGAGTGCATGGCGCAATCTCAGCTACCGC 74
QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
DB 75 AACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCTCAGCCTCCCGCAGTAGCTGGATT 134
QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrrPhePheLeuValGluMetGluPhe 60
DB 135 ACAGGATGTGCACCCAGCTCGGCTAAATTTTGTATTTTATTTAGTAGAGTGGATT 194
QY 61 LeuHisValGlyGlnAlaGlyLeuProThrSerAspSerProSerValSerAla 80
DB 195 CTCATGTGGTCAGGCTGTCTCGAATCTCCGACCTCAGATGATCCCTCGTCTCGGCC 254
QY 81 SerGlnSerAlaArgTyrrArgThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCys 100
DB 255 TCCCAAGTGTAGATACAGGACTGGCCACCATGCGCGCTCTGCTGCTAAATTTTGT 314
QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120
DB 315 GSTAGAAACAGGGTTTTCACGTATGTCCTCCCAAGCTGCTCTCCTAGCTCAAGCAGTCCACC 374
QY 121 CysLeuSerLeuProLysCysTyrAspTyrArgAlaAlaValProGlyLeuPheIle 140
DB 375 TGCTCAGCTCCCAAGTGTGGATTACAGCGTGCAGCGTGCAGCGTCTGCTGCTTTTAT 434
QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
DB 435 TTATTTTAAAGACACAGGTGTCCTCCACTCTTTACCCAGGATGAAGTGCAGTGGTGTGAT 494
QY 161 HisSerSerLeuGlnProSerThrProGluIleLysHisProProAlaSerAlaSerGln 180
DB 495 CACAGCTCACTGCAGCGCTTCACTCTCTGAGATCAAGCATCTCTGCTCAGCCTCCCAA 554
QY 181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIlePheAsn 200
DB 555 GTAGCTGGGACCAAGACATGCACCATCAGCTGGCTAAATTTTATTTTATTTTAT 614
QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
DB 615 TTTTTCAGACAGAGTCTCAACTCTGTCCACCCAGCTGGAGTGCAGTGGCGCAATCTTGGC 674
QY 221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
DB 675 TCACGTCAACTCTGCTCCCGGTTCAAGTTATTTCTCTGCCCGCAGCTCTCTAGTAGC 734
QY 241 TrpAspTyrArgArgProProArgLeuAlaAsnPheValPheLeuValGluMetGly 260
DB 735 TGGAGCTACAGGCGCCACCCAGCCTAGCTAAATTTTGTATTTTATTTAGTAGAGTGGG 794
QY 261 PheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeuProAlaSerAla 280
```

```
QY 81 SerGlnSerAlaArgTyrArgThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCys 100
Db 255 TCCCAAAGTGTAGATACAGGACTGGCCACCACATGCCCGCTCTGCCTGGCTAATTTTGT 314
QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120
Db 315 GGTAGAAAACAGGGTTTCACTGATGTGCCCAAGCTGCTCTCCTGAGCTCAAGCAGTCCACC 374
QY 121 CysLeuSerLeuProLysCysTrpAspTyrArgAlaAlaValProGlyLeuPheile 140
Db 375 TGCCTCAGCCTCCCAAAGTGTGGATTACAGCGGTGCAGCGTGCCTGGCTGCTTTTATT 434
QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
Db 435 TTATTTTTTTAAGACACAGGTGTCCCACTCTTACCAGGATGAAGTGCAGTGGTGTGAT 494
QY 161 HisSerSerLeuGlnProSerThrProGluLeuLysHisProAlaSerAlaSerGln 180
Db 495 CACAGCTCACTCAGCGCTTCAACTCTCTGAGATCAAGCATCTCTCCTCAGCTCCCA 554
QY 181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePhePheilePheAsn 200
Db 555 GTAGCTGGGACCAAGACATGACCACCTACACTGCTGCTAATTTTATTATTTTAAAT 614
QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db 615 TTTTGTAGACAGAGTCTCAACTCTGTACCCAGGTGGAGTGCAGTGGCGCAATCTTGGC 674
QY 221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
Db 675 TCACATGCACTCTGCTCCCGGTTCAAGTTATTTCTCTGCCACAGCTCTCTGAGTAGC 734
QY 241 TrpAspTyrArgArgProProArgLeuAlaAsnPheValPheLeuValGluMetGly 260
Db 735 TGGGACTACAGCGCGCCACACAGCGCTAGCTAATTTTTTGTATTTTGTAGAGATGGGG 794
QY 261 PheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeuProAlaSerAla 280
Db 795 TTCACCATGTTCCGACAGGTGATCTTGATCTCTGGACCTTGATCTGCCTGCCCTGGGCC 854
QY 281 SerGlnSerAlaGlyIleThrGlyValSerHisHisAlaArgLeuIlePheAsnPheCys 300
Db 855 TCCCAAAGTGTGGGATTACAGCGGTGAGCCACCGCCCGCTTATTTTAAATTTTGT 914
QY 301 LeuPheGluMetGlnSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320
Db 915 TTGTTTGAATGGAATCTCACTCTGTATCCAGGCTGGAGTGAATGGCCAAATCTCGGC 974
QY 321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340
Db 975 TCACTGCAACCTCTGCCTCCCGGCTCAAGCGATTCTCTGTCTCAGCCTCCCAAGCAGC 1034
QY 341 TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheileArgGlyGly 360
Db 1035 TGGGATTACGGCACCTGCCACACACACCCCGCTAATTTTGTATTTTCATTAGAGCGGG 1094
QY 361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375
Db 1095 GTTTCACCATATTTGTAGCGCTGGTCTCAAACTCTCAGCTCAGCTCAGG 1139
```

## RESULT 9

US-10-198-069-48

; Sequence 48, Application US/10198069

; Publication No. US20030096756A1

; GENERAL INFORMATION:

; APPLICANT: AVERBACK, PAUL

; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER

; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF

; TITLE OF INVENTION: CELLS

; FILE REFERENCE: 59003.000009

; CURRENT APPLICATION NUMBER: US/10/198,069

; CURRENT FILING DATE: 2002-07-19

; PRIOR APPLICATION NUMBER: 60/306,161

```
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/306,150
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/331,477
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 1442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15)..(1139)
US-10-198-069-48

Alignment Scores:
Pred. No.: 7,83e-197 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-380-203-2 (1-375) x US-10-198-069-48 (1-1442)
QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
Db 15 ATGGAGTTTTCGCTCTTGTGCCAGGCTGGAGTGCATATGGCGCAATCTCAGCTCACC GC 74
QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
Db 75 AACCTCCCGCTCCCGGGTTCAAGCGATTCTCTGCCTCAGCCTCCCGTAGTAGTGGGATT 134
QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60
Db 135 ACAGCATGTGCACCCACGCTCGGCTAATTTTGTATTTTATTAGTAGAGATGGAGTTT 194
QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspProSerValSerAla 80
Db 195 CTCCTATGTGTGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATCCCTCCGCTCGGCC 254
QY 81 SerGlnSerAlaArgTyrArgThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCys 100
Db 255 TCCCAAAGTGTGATAGATACAGGACTGGCCACCATGCGCTGCTGCCTGGCTAATTTTGT 314
QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120
Db 315 GGTAGAAAACAGGGTTTCACTGATGTGCCCAAGCTGCTCTCTGAGCTCAAGCAGTCCACC 374
QY 121 CysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheile 140
Db 375 TGCCTCAGCCTCCCAAAGTGTGGATTACAGCGGTGCAGCGTGCCTGGCTGCTTTTATT 434
QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
Db 435 TTATTTTTTTAAGACACAGGTGTCCCACTCTTACCAGGATGAAGTGCAGTGGTGTGAT 494
QY 161 HisSerSerLeuGlnProSerThrProGluLeuLysHisProAlaSerAlaSerGln 180
Db 495 CACAGCTCACTCAGCGCTTCAACTCTCTGAGATCAAGCATCTCTCCTCAGCTCCCA 554
QY 181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePhePheilePheAsn 200
Db 555 GTAGCTGGGACCAAGACATGACCACCTACACTGCTGCTAATTTTATTATTTTAAAT 614
QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db 615 TTTTGTAGACAGAGTCTCAACTCTGTCCAGGCTGGAGTGCAGTGGCGCAATCTTGGC 674
QY 221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
Db 675 TCACATGCACTCTGCTCCCGGTTCAAGTTATTTCTCTGCCACAGCTCTCTGAGTAGC 734
```



```
US-10-755-889-409
; Sequence 409, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; PRIOR FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 409
; LENGTH: 1442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-755-889-409

Alignment Scores:
Pred. No.: 7,83e-197 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-09-380-203-2 (1-375) x US-10-755-889-409 (1-1442)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
Db 15 ATGGAGTTTTCGCTCTTGTGGCGGTTCAAGCGATCTCTGCTCAGCGTCCCGCAATCTCAGCTCACCOC 74

QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValaAlaGlyIle 40
Db 75 AACCTCCGCTCCCGGGTTCAAGCGATCTCTGCTCAGCGTCCCGCAATCTCAGCTCAGGATT 134

QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60
Db 135 ACAGGCGATGTCACCCAGCTCGGCTAATTTTGTATTTTGTAGTAGAGATGGAGTTT 194

QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspSerProSerValSerAla 80
Db 195 CTCCTATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATCCCTCGCTCGGCC 254

QY 81 SerGlnSerAlaArgTyrArgThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCys 100
Db 255 TCCCAAAGTGTAGATACAGGACTGGCCACCATGCCCGCTCTGCTGGCTAATTTTGT 314

QY 101 GlyArgAsnArgValSerLeuMetCysProSerTyrSerProGluLeuLysGlnSerThr 120
Db 315 GGTAGAACAGGGTTTCACTGATGTGCCCAAGTGGTCTCTCCTGAGTCAAGCAGTCCACC 374

QY 121 CysLeuSerLeuProLysCysTyrAspTyrArgArgAlaAlaValProGlyLeuPheIle 140
Db 375 TGCCTCAGCTCCCAAGTGTGGGATTACAGCGGTGACCGCTGCTGCTGCTGCTGCTGCTGCT 434

QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTyrCysAsp 160
Db 435 TTATTTTTTTTAAACACAGGTGTCCACTCTTACCAGGATGAGTGGAGTGGTGGTGGT 494

QY 161 HisSerSerLeuGlnProSerThrProGluIleLysHisProProAlaSerAlaSerGln 180
Db 495 CACAGCTCACTCAGCGCTTCAACTCTGAGATCAAGCATCTCTGCTCAGCGTCCCA 554

QY 181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIlePheIlePheAsn 200
Db 555 GTAGCTGGGACCAAGACATGACCACTACACTGCTGCTAATTTTATTTTATTTTAT 614

QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db 75 AACCTCCGCTCCCGGGTTCAAGCGATCTCTGCTCAGCGTCCCGCAATCTCAGCTCACCOC 74

US-10-755-889-409
; Sequence 409, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; PRIOR FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 409
; LENGTH: 1442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-755-889-409

Alignment Scores:
Pred. No.: 7,83e-197 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-10-910-173-1
; Sequence 1, Application US/10910173
; Publication No. US20050090441A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R
; APPLICANT: de la Monte, Suzanne M
; TITLE OF INVENTION: Inhibition of Neurodegeneration
; FILE REFERENCE: 21486-047
; CURRENT APPLICATION NUMBER: US/10/910,173
; CURRENT FILING DATE: 2004-08-02
; PRIOR APPLICATION NUMBER: US/09/872,968
; PRIOR FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-910-173-1

Alignment Scores:
Pred. No.: 7,83e-197 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-09-380-203-2 (1-375) x US-10-910-173-1 (1-1442)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
Db 15 ATGGAGTTTTCGCTCTTGTGGCGGTTCAAGCGATCTCTGCTCAGCGTCCCGCAATCTCAGCTCACCOC 74

QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValaAlaGlyIle 40
Db 75 AACCTCCGCTCCCGGGTTCAAGCGATCTCTGCTCAGCGTCCCGCAATCTCAGCTCACCOC 74
```





QY 120 hrCysLeuSerLeuProLysCysTrpAspTyrArgArgAlaValProGlyLeuPheI 140  
Db 372 CTGGCTCAGCTCCCAAGTGTGGATTACAGCGTGCACCGCGCTGGCCCTTTTAA 431  
QY 140 leLeuPheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysA 160  
Db 432 TTTTATTTTAAAGACACAGGTGTCCCACTCTTACCAGGATGAAGTGCAGTGGTG 491  
QY 160 spHisSerLeuGlnProSerThrProGluLeuLysHisProProAlaSerAlaSerG 180  
Db 492 ATCACAGTCTCACTGACGCTTCAACTCT-GAGATCAAGCATCTCTCGCTCAGCCTCC 550  
QY 180 ln-ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuPhePhelePhe 199  
Db 551 AAAGTAGCTGGGACCAAGACATGACCACTACACTGCTGCTAAATTTTATTTTATTTT 610  
QY 200 AnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAnLeu 219  
Db 611 AATTTTGTGACACAGTCTCACTCTGTACCCAGGCTGGAGTGCAGTGGCGCAATCTT 670  
QY 220 GlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSer 239  
Db 671 GGCTCACTGCAACCTCTGCTCCCGGTTCAAGTTATTCTCTGCTGCCAGCCTCCTGAGT 730  
QY 240 SerTrpAspTyrArgArgProProArgLeuAlaAsnPheValPheLeuValGluMet 259  
Db 731 AGCTGGGACTACAGGCGCCACACAGCTGTACCCAGGCTGGAGTGCAGTGGCGCAATCTT 790  
QY 260 -GlyPheThrMetPheAlaArgLeuLeu- IleSerGlyProCysAspLeuProAlaS 279  
Db 791 GGGTTTCAACATGTTCCGACGTTGATGCTAGATCTCTTGACCTTGTATCTGCTGCT 850  
QY 279 erAlaSerGlnSerAlaGlyIleThrGlyValSerHisHis-AlaArgLeuLeuPheAsn 298  
Db 851 CGGCTCCCAAGTGTGGATTACAGGACGTGACGCCACCGCCGCGCTATTTTAAAT 910  
QY 299 PheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsn 318  
Db 911 TTTTGTGTTGTAATGAATGAATCTCACTGTGTACCCAGGCTGGAGTGCAGTGGCGCAAT 970  
QY 319 LeuGlySerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuPro 338  
Db 971 CTCGGCTCACTGCAACCTCTGCTCCCGGCTCAAGCGATTCTCTGTCTCAGCCTCCCA 1030  
QY 339 SerSerTrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheleArg 358  
Db 1031 AGCAGCTGGGATTACGGGCACCTGCA-CCACACCCCGCTAAATTTTGTATTTTCAATTAGA 1089  
QY 359 GlyGlyValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375  
Db 1090 GCGGGGTTTACCATAATTTGTGAGGCTGGTCTCAAACTCTCGACCTCAGG 1140

## RESULT 14

US-09-964-412-4  
; Sequence 4, Application US/09964412  
; Patent No. US20020129391A1  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; Wands, Jack R.  
; TITLE OF INVENTION: Transgenic Animals and Cell Lines for  
; Screening Drugs Effective for the Treatment or Prevention  
; of Alzheimer's Disease  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.  
; STREET: 1100 New York Ave., Suite 600  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/964,412  
FILING DATE: 28-Sep-2001  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0609.4370000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1418 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-964-412-4  
Alignment Scores:  
Pred. No.: 1,08e-171 Length: 1418  
Score: 1787.00 Matches: 363  
Percent Similarity: 95.54% Conservative: 1  
Best Local Similarity: 95.28% Mismatches: 11  
Query Match: 87.86% Indels: 10  
DB: Gaps: 0  
US-09-380-203-2 (1-375) x US-09-964-412-4 (1-1418)  
QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20  
Db 14 ATGGAGTTTTCGCTCTTGTTCGCCAGGTGAGTGCATGCGCAATCTCAGCTCACC 73  
QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProVal-AlaGlyI 40  
Db 74 AACCTCCGCTCCCGGTTCAAGCATTTCTCTGCTCAGCCTCCCAAGTAGGCTGGAT 133  
QY 40 eThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPh 60  
Db 134 TACAGGACATGTCAC-CACGCTCGGCTAAATTTTGTATTTTATTTAGTAGAGTAGG 192  
QY 60 eLeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspSerProSerValSerAl 80  
Db 193 TCTCATGTTGGTTCAGGCTGGTCTCGAACT-CCGACCTCAGATGATCTCCCGTCTCG 251  
QY 80 aSerGlnSerAlaArgTyrArgThrGlyHisHisAlaArg-LeuCysLeuAlaAsnPheC 100  
Db 252 CTCCCAAGTCTAGATACAGGACTGAGCACCATGCGCGGCTCTGCTGGCTAAATTTT 311  
QY 100 ysGlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSer 120  
Db 312 GTGGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTCCTGAGCTCAAGCAGT 371  
QY 120 hrCysLeuSerLeuProLysCysTrpAspTyrArgArgAlaValProGlyLeuPheI 140  
Db 372 CTGCTCAGCTCCCAAGTGTGGGATTACAGCGTGCAGCGGTGCTGGCCCTTTTAA 431  
QY 140 leLeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysA 160  
Db 432 TTTTATTTTAAAGACACAGGTTGCCACTCTTTACCCAGGATGAAGTGCAGTGGTG 491  
QY 160 spHisSerLeuGlnProSerThrProGluLeuLysHisProProAlaSerAlaSerG 180  
Db 492 ATCACAGTCTCACTGACGCTTCAACTCT-GAGATCAAGCATCTCTCGCTCAGCCTCC 550  
QY 180 ln-ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuPhePhelePhe 199  
Db 551 AAAGTAGCTGGGACCAAGACATGACCACTACACTGCTGCTAAATTTTATTTTATTTT 610

```

QY 200 AsnPheLeuArgGlnSerLeuAenSerValThrGlnAlaGlyValGlnTrpArgAsnLeu 219
Db 611 AATTTTGGAGACAGAGTCTCAACTGTGTCCACCCAGCGTGGAGTGGAGTGGCGCAATCTT 670
QY 220 GlySerLeuGlnProLeuProGlyPheLeuPheSerCysProSerLeuLeuSer 239
Db 671 GGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCGCCAGCTCTCTGAGT 730
QY 240 SerTrpAspTyrArgArgProArgLeuAlaAsnPhePheValPheLeuValGluMet 259
Db 731 AGCTGGAGCTACAGGGCCGCCACCGCTAGCTAAATTTTGTATTTTAGTAGAGATG 790
QY 260 -GlyPheThrMetPheAlaArgLeuLeuLeu-ileSerGlyProCysAspLeuProAlas 279
Db 791 GGGTTTTCACCATGTTCCGACAGTTGATGCTAGATCTCTTGACCTTGTGATCTGCTGCT 850
QY 279 erlAsrGlnSerAlaGlylleThrGlyValSerHisHis-AlaArgLeullePheAsn 298
Db 851 CGGCTCTCCCAAGTGTGGATTACAGGACGTGACGCCCGCCGCGCTATTTTAAAT 910
QY 299 PheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsn 318
Db 911 TTTTGTGTGTTGAAATGGAAATCTCACTCTGTTACCCAGGCTGGAGTGCATGGCCAAAT 970
QY 319 LeuGlySerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuPro 338
Db 971 CTGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCGATTCTCTGTCTCAGCTCCCA 1030
QY 339 SerSerTrpAspTyrGlyHisLeuProProHisProAlaAsnPheCyslePheleArg 358
Db 1031 AGCAGCTGGGATTACGGGACCTGCA-CCACACCCCGCTAAATTTTGTATTTTCAATGA 1089
QY 359 GlyGlyValSerProTyrLeuSerGlyTyrSerGlnThrProAspLeuArg 375
Db 1090 GCGGGGTTTACCATATTTGTGAGGTGCTGTCTCAAACTCTGACCTCAGG 1140

RESULT 15
US-09-964-667-4
; Sequence 4, Application US/09964667
; Publication No. US2003003621A1
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; Wands, Jack R.
; TITLE OF INVENTION: Transgenic Animals and Cell Lines for
; Screening Drugs Effective for the Treatment or Prevention
; Of Alzheimer's Disease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/964,667
; FILING DATE: 28-Sep-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0609.4370000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 1418 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-964-667-4

Alignment Scores:
Pred. No.: 1,08e-171 Length: 1418
Score: 1787.00 Matches: 363
Percent Similarity: 95.54% Conservatives: 1
Best Local Similarity: 95.28% Mismatches: 11
Query Match: 87.86% Indels: 10
DB: 10 Gaps: 0

US-09-380-203-2 (1-375) x US-09-964-667-4 (1-1418)
QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
Db 14 ATGGAGTTTTCGCTCTTGTGTCAGGCTGGAGTGCATGGCGCAATCTCAGCTCACCGC 73
QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProVal-AlaGlyIle 40
Db 74 AACCTCCGCTCCCGGGTTTCAAGCGATTCTCTGCTCAGCTCCCGCAGTAGCTGGAT 133
QY 40 eThrGlyMetCysThrHisAlaArgLeulleLeuTyrPhePheLeuValGluMetGluPh 60
Db 134 TACAGCATGTGCAC-CACGCTCGGCTAATTTTGTATTTTGTAGTAGAGATGAGTT 192
QY 60 eLeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAl 80
Db 193 TCTCCATGTTGTGTCAGGCTGGTCTCGAACT-CCGACCTCAGATGATCCTCCGCTCGGC 251
QY 80 aSerGlnSerAlaArgTyrArgThrGlyHisHisAlaArg-LeuCysLeuAlaAsnPheC 100
Db 252 CTCCTCAAAAGTGTAGATACAGGACTGAGCACCATTGCGCCGCTCTGCTGGCTAAATTTT 311
QY 100 ysGlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerT 120
Db 312 GTGGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTGAGCTCAAGCAGTCCA 371
QY 120 hrCysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheI 140
Db 372 CTGCTCTCAGCTCCCAAGTCTGGATTACAGCGGTGACCGGTGCGCTGGCTTTTAA 431
QY 140 leLeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysA 160
Db 432 TTTTATTTTAAAGACACAGGTCTCCACTCTTACCAGGATGAAGTGCAGTGGTGTG 491
QY 160 spHisSerLeuGlnProSerThrProGluIleIysHisProProAlaSerAlaSerG 180
Db 492 ATCAGAGTCACTGAGGCTTCAACTCT-GAGATCAAGCATCTCTCTGCTCAGCCTCCC 550
QY 180 In-ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIlePhe 199
Db 551 AAGTAGCTGGACCAAGACATGCACCACTACACTGCTGCTAAATTTTATTTATTTT 610
QY 200 AsnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeu 219
Db 611 AATTTTGGAGACAGAGTCTCAACTCTGTGTCAGGCTGGAGTGCAGTGGCGCAATCTT 670
QY 220 GlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSer 239
Db 671 GGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCGCCAGCTCTCTGAGT 730
QY 240 SerTrpAspTyrArgArgProArgLeuAlaAsnPhePheValPheLeuValGluMet 259
Db 731 AGCTGGAGCTACAGGGCCGCCACCGCTAGCTAAATTTTGTATTTTAGTAGAGATG 790
QY 260 -GlyPheThrMetPheAlaArgLeuLeuLeu-ileSerGlyProCysAspLeuProAlas 279
Db 791 GGGTTTTCACCATGTTCCGACAGTTGATGCTAGATCTCTTGACCTTGTGATCTGCTGCT 850

```

```

QY 279 erAlaSerGlnSerAlaGlyIleThrGlyValSerHisHis-AlaArgLeuIlePheAsn 298
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
851 CGGCTTCCCAAGTGTGGGATTACAGGACGTGACGCCACCGCCGCGCTATTTTAAAT 910
QY 299 PheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsn 318
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
911 TTTTGTGTTTGGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCATGGCCAAAT 970
QY 319 LeuGlySerLeuGlnProLeuProProGlyLeuLysArgPheSerCysLeuSerLeuPro 338
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
971 CTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCTGTCTCAGCCTCCCA 1030
QY 339 SerSerTrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArg 358
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1031 AGCAGCTGGGATTACGGGCACCTGCA-CCACACCCCGCTAAATTTTGTATTTTCATTAGA 1089
QY 359 GlyGlyValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1090 GCGGGGGTTTACCATTATTTGTGAGGCTGGTCTCAAACTCTGACCTCAGG 1140

```

Search completed: September 16, 2005, 01:25:05  
Job time : 828 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 15, 2005, 22:21:14 ; Search time 222 Seconds  
(without alignments)  
2763.981 Million cell updates/sec

Title: US-09-380-203-2  
Perfect score: 2034  
Sequence: 1 MFESLLPRLECNCAISAH.....FIRGGVSPVLSQWSQTPDLR 375

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09380203/runat\_13092005\_170812\_11728/app\_query.fasta\_1.519  
-DB=Issued\_Patents\_NA -QFWT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09380203 @CNC 1.1 177 @runat\_13092005\_170812\_11728 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA:  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2034	100.0	1442	2	US-08-454-557C-120
2	2034	100.0	1442	2	US-08-340-426D-120
3	2034	100.0	1442	2	US-08-450-673C-120
4	2034	100.0	1442	4	US-09-872-968-1
5	1787	87.9	1418	5	PCR-US95-17111A-120
6	1520	74.7	1381	2	US-08-454-557C-49
7	1520	74.7	1381	2	US-08-340-426D-49
8	1520	74.7	1381	2	US-08-450-673C-49
9	1520	74.7	1381	5	PCR-US95-17111A-49
10	913	44.9	9779	4	US-09-949-016-15370
11	908	44.6	10790	4	US-09-949-016-14370
12	885	43.5	50518	4	US-09-949-016-12315

13	871.5	42.8	14411	4	US-09-949-016-13718	A
14	868.5	42.7	56616	4	US-09-949-016-12462	A
15	868.5	42.7	56616	4	US-09-949-016-17085	A
16	866	42.6	60424	4	US-09-949-016-12175	A
17	849.5	41.8	100463	4	US-09-949-016-12511	A
18	849.5	41.8	100468	4	US-09-949-016-13725	A
19	845	41.5	28556	4	US-09-949-016-13064	A
20	845	41.5	40037	4	US-09-949-016-12715	A
21	844	41.5	26104	4	US-09-949-016-14045	A
22	844	41.5	77626	4	US-09-949-016-12608	A
23	840	41.3	60489	4	US-09-949-016-16287	A
24	838	41.2	18291	4	US-09-949-016-14787	A
25	838	41.2	36952	4	US-09-949-016-14786	A
26	838	41.2	100463	4	US-09-949-016-12511	A
27	838	41.2	100468	4	US-09-949-016-13725	A
28	837.5	41.2	36311	4	US-09-949-016-13627	A
29	837	41.2	14796	3	US-08-975-080-35	Appl
30	837	41.2	14796	3	US-09-630-706-10	Appl
31	837	41.2	14796	3	US-09-496-694B-3	Appl
32	837	41.2	14796	4	US-10-138-618-35	Appl
33	837	41.2	14796	4	US-09-918-186A-3	Appl
34	833.5	41.0	16063	3	US-09-801-052-3	Appl
35	833.5	41.0	16063	4	US-10-020-121-3	Appl
36	832	40.9	42693	4	US-09-949-016-17317	A
37	832	40.9	42693	4	US-09-949-016-17318	A
38	828.5	40.7	25656	4	US-09-949-016-13022	A
39	828	40.7	17198	4	US-09-949-016-13776	A
40	827	40.7	15202	4	US-09-949-016-16883	A
41	825	40.6	27525	4	US-09-949-016-16835	A
42	824	40.5	15585	4	US-09-949-016-11927	A
43	824	40.5	15585	4	US-09-949-016-15627	A
44	824	40.5	50950	4	US-09-949-016-16659	A
45	823.5	40.5	34725	4	US-09-949-016-15797	A

## ALIGNMENTS

RESULT 1  
US-08-454-557C-120  
; Sequence 120, Application US/08454557C  
; Patent No. 5830670  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 30-MAY-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ludwig, Steven R.  
; REGISTRATION NUMBER: 36,203  
; REFERENCE/DOCKET NUMBER: 0609.3840003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 120:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1442 base pairs

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..1139
US-08-454-557C-120

Alignment Scores:
Pred. No.: 3,52e-184 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 2

US-09-380-203-2 (1-375) x US-08-454-557C-120 (1-1442)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
Db 15 ATGGAGTTTTCGCTCTTGTTCAGGCTGGAGTGCATGGGCAATCTCAGCTCAGCCG 74
QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
Db 75 AACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCTCAGCTCCCGAGTAGCTGGATT 134
QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60
Db 135 ACAGGATGTCACCCACGCTCGGCTAATTTGTATTTTGTATTTTGTAGTAGAGTTT 194
QY 61 LeuHisValGlyClnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAla 80
Db 195 CTCATGTTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCCCTCCGCTCGGCC 254
QY 81 SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys 100
Db 255 TCCCAAGTGTAGATACAGGACTGGCCACCATGCCCCGCTCTGCTGGCTAATTTTGT 314
QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120
Db 315 GGTAGAACAGGGTTTCACTGATGTGCCCAAGCTGCTCTCTGAGCTCAAGCAGTCCACC 374
QY 121 CysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheIle 140
Db 375 TGCTTCAGCTCCCAAGTGTGGGATTACAGCGTGCAGCGTGCCTCGCCCTTTTATT 434
QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
Db 435 TTATTTTTTTAAGACACAGGTGTCCCACTCTTACCAGGATGAAGTGCAGTGGTGTAT 494
QY 161 HisSerSerLeuGlnProSerThrProGluIleLysHisProProAlaSerAlaSerGln 180
Db 495 CACAGCTCACTGCAGCTTCAACTCTGAGATCAAGCATCTCTCTGCTCAGCTCCCAA 554
QY 181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIlePheAsn 200
Db 555 GTAGCTGGGACCAAGACATGCACCATACCTGAGTCAAGCATCTCTGCTCAGCTCCCAA 614
QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db 615 TTTTGTAGACAGAGTCTCAACTCTGTACCCAGGCTGGAGTGCAGTGGCGCAATCTTGGC 674
QY 221 SerLeuGlnProLeuProProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
Db 675 TCACTGCAACCTCTGCTCCCGGTTCAAGTTATTCTCTGCCCGCAGCTCTCTGAGTAGC 734
QY 241 TrpAspTyrArgArgProArgLeuAlaAsnPhePheValPheLeuValGluMetGly 260
Db 735 TGGGACTACAGGCGCCACACGCCCTAGTAAATTTTGTATTTTGTAGAGATGGGG 794
QY 261 PheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeuProAlaSerAla 280

```

```

Db 795 TTCACATGTTCCGCCAGGTTGATCTTGATCTCTGGACCTTGATCTGCCTCGGCC 854
QY 281 SerGlnSerAlaGlyIleThrGlyValSerHisHisAlaArgLeuIlePheAsnPheCys 300
Db 855 TCCCAAGTGTCTGGATTACAGCGTGCAGCCACGCGCGCTATTATTTAATTTTGT 914
QY 301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320
Db 915 TTGTTTGAATGGAATCTCACTCTGTATCCCGAGCTGGAGTGCATGGCCAAATCTCGGC 974
QY 321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340
Db 975 TCACGTGCAACCTCTCTCCCTCCCGGCTCAAGCATTTCTCTGCTCAGCTCCCAAGCAGC 1034
QY 341 TrpAspTyrGlyHisLeuProHisProAlaAsnPheCysIlePheIleArgGlyGly 360
Db 1035 TGGGATTACGGGCACCTGCCACACACCCCGCTAATTTTGTATTTTCAATTAGAGCGGG 1094
QY 361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375
Db 1095 GTTTCACCATATTTTGTGAGGCTGGTCTCAAACTCTCGACCTCAGG 1139

RESULT 2
US-08-340-426D-120
; Sequence 120, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 120:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1442 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..1139
US-08-340-426D-120

Alignment Scores:
Pred. No.: 3,52e-184 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

```

DB:	2	Gaps:	0
US-09-380-203-2	(1-375) x US-08-340-426D-120	(1-1442)	
Qy	1	MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg	20
Db	15	ATGGAGTTTTCGCTCTTGTGTGCCAGGCTGGAGTGCATATGGCGCAATCTCAGCTCACCGC	74
Qy	21	AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle	40
Db	75	AACCTCGCCTCCCGGTTCAAGCGATTCTCTCGCTCAGCCTCCCCAGTAGCTGGGATT	134
Qy	41	ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe	60
Db	135	ACAGGCATGTGCACCCACGCTCGGCTAAATTTGTATTATTTTATTTTATTTAGTAGAGATGAGATT	194
Qy	61	LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspProSerValSerAla	80
Db	195	CTCCATGTGTGAGCTGGTCTCGAACTCCCGACCTCAGATGATCCTTCGCTCGGCC	254
Qy	81	SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys	100
Db	255	TCCCAAGTGTAGATACAGGACTGGCCACCATGCCCGGCTCTCGCTGGCTAAATTTTGT	314
Qy	101	GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr	120
Db	315	GGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTCGAGCTCAAGCAGTCCAC	374
Qy	121	CysLeuSerLeuProLysCysTyrAspTyrArgArgAlaAlaValProGlyLeuPheIle	140
Db	375	TGCCTCAGCCTCCCAAGTGTGGATTCAGAGCGTGCAGCGGTGCTCGCCTTTTATT	434
Qy	141	LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp	160
Db	435	TTATTTTAAAGACACAGGTGTCACCTCTTACCCAGGATGAAGTGCAGTGGTGTGAT	494
Qy	161	HisSerSerLeuGlnProSerThrProGluIleLysHisProProAlaSerAlaSerGln	180
Db	495	CACAGCTCACTGCAGCCTTCAACTCTCTGAGATCAAGCATCTCTCGCTCAGCCTCCCAA	554
Qy	181	ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIlePheAsn	200
Db	555	GTAGCTGGGACCAAGACATGCACCTACACCTGGCTAAATTTTATTTTATTTTAAAT	614
Qy	201	PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly	220
Db	615	TTTTTGTAGACAGAGTCTCAACTCTGTCCACCCAGGCTGGAGTGCAGTGGCGCAATCTTGGC	674
Qy	221	SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer	240
Db	675	TCAGTCAAACTCTGCCTCCCGGTTCAAGTTATTCTCTCGCCACGCTCCTGAGTAGC	734
Qy	241	TrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeuValGluMetGly	260
Db	735	TGGGACTACAGGCGCCACACGCTAGCTAAATTTTGTGATTTTGTAGTAGAGATGGG	794
Qy	261	PheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeuProAlaSerAla	280
Db	795	TTACCAATGTGGCCAGGTTGATCTTGATCTCTGGAACCTTGATCTCGCCTCGGCC	854
Qy	281	SerGlnSerAlaGlyIleThrGlyValSerHisAlaArgLeuIlePheAsnPheCys	300
Db	855	TCCCAAGTGTGGGATTACAGGCGTGCAGCCACCCAGCGCTTATTTTAAATTTTGT	914
Qy	301	LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly	320
Db	915	TTGTTTGAATAATCTCACTCTGTGTACCCAGGCTGGAGTGCATGGCCAAATCTCGGC	974
Qy	321	SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer	340
Db	975	TCAGTCAAACTCTGCCTCCCGGCTCAAGGATTCTCTGTCTCAGCCTCCCAAGCAGC	1034
Qy	341	TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGlyGly	360

```

Db 1035 TGGGATTACGGGACCTGCCACCACACCCCGCTAAATTTTGTATTTTCATTAGAGCGGG 1094
Qy 361 ValSerProTyRLeuSerGlyTyrSerGlnThrProAspLeuArg 375
Db 1095 GTTTCACCATATTGTGACGGCTGGTCTCAAACTCTGACCTCAGG 1139

RESULT 3
US-08-450-673C-120
; Sequence 120, Application US/08450673C
; Patent No. 594888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 120:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1442 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 15..1139
US-08-450-673C-120

Alignment Scores:
Pred. No.: 3,52e-184 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-380-203-2 (1-375) x US-08-450-673C-120 (1-1442)

Qy 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
Db 15 ATGAGATTTTCGCTCTTTGTCGCCAGCTGGAGTGCAATGGCCCAATCTCAGCTCACCGC 74
Qy 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValaIleGlyIle 40
Db 75 AACCTCGCGCTCCGGGTTCAAGCGATTCTCTCGCTCAGCCTCCCGAGTAGCTGGGATT 134
Qy 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyPhePheLeuValGluMetGluPhe 60
Db 135 ACAGGATGTGACACCCACCGCTCGGCTAAATTTGTATTTTATTTAGTAGAGATGGAGTTT 194

```

```

QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAla 80
Db 195 CTCACATGTTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCCCTCCGCTCGGCC 254

QY 81 SerGlnSerAlaArgTyrArgThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCys 100
Db 255 TCCCAAAGTGTAGATACAGGACTGCCACCAATGCCGGCTCTGGCTGGCTAAATTTTGT 314

QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120
Db 315 GGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGCTCTCTGAGCTCAAGCAGTCCACC 374

QY 121 CysLeuSerLeuProLysCysTrpAspTyrArgAlaAlaValProGlyLeuPheLe 140
Db 375 TGCCTCAGCCTCCCAAGTGTGGGATACAGCGGTGCAGCGTGCAGCGTCTGGCTTTTAT 434

QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
Db 435 TTATTTTAAAGACACAGGTGTCCCACTCTTACCCAGATGAAGTGCAGTGGTGTGAT 494

QY 161 HisSerSerLeuGlnProSerThrProGluLeuLysHisProProAlaSerAlaSerGln 180
Db 495 CACAGCTCACTCAGCGCTTCAACTCTCTGAGATCAAGCATCTCTCTGCTCAGCCTCCCA 554

QY 181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuLeuPhePheLePheAsn 200
Db 555 GTAGCTGGGACCAAGACATGCACCATACACCTGCTAAATTTTATTTTATTTTAAAT 614

QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db 615 TTTTGTGAGACAGAGTCTCAACTCTGTACCCAGGCTGGAGTGCATGGCGCAATCTTGGC 674

QY 221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
Db 675 TCACGTCAACCTCTGCTCCCGGTTCAAGTTATTTCTCTGCCCGCAGCTCTCTGAGTAGC 734

QY 241 TrpAspTyrArgProProArgLeuAlaAsnPheValPheLeuValGluMetGly 260
Db 735 TGGGACTACAGCGGCCACACAGCGCTAGCTAAATTTTGTATTTTATTTAGTAGAGTGGG 794

QY 261 PheThrMetPheAlaArgLeuLeuLeuSerGlyProCysAspLeuProAlaSerAla 280
Db 795 TTCACATGTTGCCAGGTTGATCTTGATCTCTGACCTTGATCTGCTCCCTCGGCC 854

QY 281 SerGlnSerAlaGlyIleThrGlyValSerHisHisAlaArgLeuLeuPheAsnPheCys 300
Db 855 TCCCAAAGTGTGGGATTACAGCGGTGAGCCACCAAGCTGCTTATTTTAAATTTTGT 914

QY 301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320
Db 915 TTGTTTGAATGGAAATCTCACTCTGTATCCCAAGGCTGGAGTGCATGGCCAAATCTCGC 974

QY 321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340
Db 975 TCACGTGCAACCTCTGCTCCCGGCTCAAGCATCTCTGCTCTGAGCTCCCAAGCAGC 1034

QY 341 TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheLeuArgGlyGly 360
Db 1035 TGGGATTACGGGCACCTGCCACCAACCCCGCTAAATTTTGTATTTTATTTAGAGCGGG 1094

QY 361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375
Db 1095 GTTTCACCATATTTGTACGGCTGGTCTCAAACTCTCTGACCTCAGG 1139

```

RESULT 4

```

US-09-872-968-1
; Sequence 1, Application US/09872968
; Patent No. 6770797
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R
; APPLICANT: de la Monte, Suzanne M
; TITLE OF INVENTION: Inhibition of Neurodegeneration

```

```

; FILE REFERENCE: 21486-047
; CURRENT APPLICATION NUMBER: US/09/872,968
; CURRENT FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-872-968-1

Alignment Scores:
Pred. No.: 3,52e-184 Length: 1442
Score: 2034.00 Matches: 375
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

```

US-09-380-203-2 (1-375) x US-09-872-968-1 (1-1442)

```

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
Db 15 ATGGAGTTTTCGCTCTTGTGGCCAGGCTGGAGTGCATGGCGCAATCTCAGCTCACCGC 74

QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
Db 75 AACCTCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGCGCTCCCGCAGTAGCTGGGATT 134

QY 41 ThrGlyMetCysThrHisAlaArgLeuLeuLeuTyrPhePheLeuValGluMetGluPhe 60
Db 135 ACAGCATGTGCACCCAGCTCGGCTAAATTTTATTTTATTTAGTAGAGTAGAGTTT 194

QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAla 80
Db 195 CTCACATGTTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCCCTCCGCTCGGCC 254

QY 81 SerGlnSerAlaArgTyrArgThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCys 100
Db 255 TCCCAAAGTGTAGATACAGGACTGGCCACCATGCGCGCTCTGCTGGCTAAATTTTGT 314

QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120
Db 315 GGTAGAAACAGGGTTTCACTGATGTGCCAAGTGTCTCTGAGCTCAAGCAGTCCACC 374

QY 121 CysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheIle 140
Db 375 TGCCTCAGCCTCCCAAGTGTGGGATTACAGCGGTGCAGCGCTGCGCTTTTATTTAT 434

QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160
Db 435 TTATTTTAAAGACACAGGTGTCCCACTCTGAGATCAAGCATCTCTGCTCAGCCTCCCA 494

QY 161 HisSerSerLeuGlnProSerThrProGluLeuLysHisProProAlaSerAlaSerGln 180
Db 495 CACAGCTCACTCAGCGCTTCAACTCTCTGAGATCAAGCATCTCTGCTCAGCCTCCCA 554

QY 181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuLeuPheIlePheLeuPheAsn 200
Db 555 GTAGCTGGGACCAAGACATGCACCATACACCTGCTAAATTTTATTTTATTTTAAAT 614

QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220
Db 615 TTTTGTGAGACAGAGTCTCAACTCTGTACCCAGGCTGGAGTGCATGGCGCAATCTTGGC 674

QY 221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240
Db 675 TCACGTGCAACCTCTGCTCCCGGTTCAAGTTATTTCTCTGCCCGCAGCTCTCTGAGTAGC 734

QY 241 TrpAspTyrArgProProArgLeuAlaAsnPheValPheLeuValGluMetGly 260
Db 735 TGGGACTACAGCGGCCACACAGCGCTAGCTAAATTTTGTATTTTATTTAGTAGAGTGGG 794

```



QY 261 PheThrMetPheAlaArgLeuLeuLeuSerGlyProCysAspLeuProAlaSerAla 280  
 Db 795 TTCACCATGTTGCCAGGTTGATCTTGACCTCTGGACCTTGATCTGCTCGGCC 854  
 QY 281 SerGlnSerAlaGlyLeuThrGlyValSerHisHisAlaArgLeuLeuPheAsnPheCys 300  
 Db 855 TCCCAAAGTCTGGGATACAGCGGTGAGCACCGCCGGCTTATTTTAAATTTTGT 914  
 QY 301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320  
 Db 915 TTGTTTGAATGGAATCTCACTCTGTACCAGGCTGGAGTGAATGGCCAAATCTCGGC 974  
 QY 321 SerLeuGlnProLeuProProGlyLeuLeuArgPheSerCysLeuSerLeuProSerSer 340  
 Db 975 TCACGTCAACCTCTGCTCCCGGCTCAAGCGATTCTCTGTCTCAGCCTCCCAAGCAGC 1034  
 QY 341 TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGlyGly 360  
 Db 1035 TGGGATACGGGACCTGCCACACACCCCGCTAAATTTTGTATTTTCAATTAGAGCGGG 1094  
 QY 361 ValSerProTyrLeuSerGlyTyrSerGlnThrProAspLeuArg 375  
 Db 1095 GTTTCACCATATTGTGACGCTGGTCTCAAACTCTTGACCTCAGG 1139

## RESULT 5

PCT-US95-17111A-120  
 ; Sequence 120, Application PC/TUS9517111A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: de la Monte, Suzanne  
 ; TITLE OF INVENTION: Neural Thread Protein Gene Expression and  
 ; TITLE OF INVENTION: Detection of Alzheimer's Disease  
 ; NUMBER OF SEQUENCES: 121  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
 ; STREET: 1100 New York Avenue, Suite 600  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20005-3934  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/17111A  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/340,426  
 ; FILING DATE: 14-NOV-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Ludwig, Steven R.  
 ; REGISTRATION NUMBER: 36,203  
 ; REFERENCE/DOCKET NUMBER: 0609.3840002  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 371-2600  
 ; TELEFAX: (202) 371-2540  
 ; INFORMATION FOR SEQ ID NO: 120:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1418 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: both  
 ; TOPOLOGY: both  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 14..1418  
 PCT-US95-17111A-120

## Alignment Scores:

Pred. No.: 1.2e-160 Length: 1418  
 Score: 1787.00 Matches: 363

Percent Similarity: 95.54% Conservative: 1  
 Best Local Similarity: 95.28% Mismatches: 11  
 Query Match: 87.86% Indels: 10  
 DB: 5 Gaps: 0  
 US-09-380-203-2 (1-375) x PCT-US95-17111A-120 (1-1418)  
 QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20  
 Db 14 ATGGAGTTTTCGCTCTTGTGGCCAGGCTGGAGTGCATGGCGCAATCTCAGCTCACCGC 73  
 QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProVal-AlaGlyIle 40  
 Db 74 AACCTCCGGCTCCGGGTTCAAGCGATTCTCTGCTCAGCCTCCCGCAGTAGGCTGGAT 133  
 QY 40 eThrGlyMetCysThrHisAlaArgLeuLeuLeuTyrPhePheLeuValGluMetGluPh 60  
 Db 134 TACAGCATGTGCAC-CACGCTCGGCTAAATTTGTATTTTTTTTAGTAGAGATGAGATT 192  
 QY 60 eLeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspSerProSerValSerAl 80  
 Db 193 TCTCCATGTTGTCAGGCTGGTCTCGAACT-CGACCTCAGATGATCTCTCCGCTCTCGGC 251  
 QY 80 aSerGlnSerAlaArgTyrArgThrGlyHisAlaArg-LeuCysLeuAlaAsnPheC 100  
 Db 252 CTCCCAAAGTGTAGATACAGACTGAGCACCATGCCCGCCTCTGCTGGCTAAATTTT 311  
 QY 100 yGlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuValGlnSerTr 120  
 Db 312 GTGGTAGAAACAGGGTTTCACTGATGTGCCAAGTGGTCTCTCTGAGCTCAAGCAGTCCA 371  
 QY 120 hrCysLeuSerLeuProLysCysTyrAspTyrArgArgAlaAlaValProGlyLeuPheI 140  
 Db 372 CTGCTCAGCCTCCCAAAGTCTGGGATTACAGGCTGCAGCGCTGCTGCGCTTTT 431  
 QY 140 leLeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysA 160  
 Db 432 TTTTATTTTTTAAAGACACAGGTGTCCCACTTTACCAGGATGAAGTGCAGTGGTGTG 491  
 QY 160 sPHisSerSerLeuGlnProSerThrProGluLeuIleHisProProAlaSerAlaSerG 180  
 Db 492 ATCAGAGCTCACTGCAGCCTTCAACTCT-GAGATCAAGCATCTCTCTGCTCAGCCTCCC 550  
 QY 180 In-ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIlePhe 199  
 Db 551 AAAGTAGCTGGGACCAAGACATGCACCTACACTACCTGGCTAAATTTTATTTATTTT 610  
 QY 200 AsnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeu 219  
 Db 611 AATTTTTTGGAGACAGAGTCTCAACTCTGTCAACCGGCTGGAGTGCAGTGGCGCAATCTT 670  
 QY 220 GlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSer 239  
 Db 671 GGCTCACTGCAACCTCTGCTCCCGGTTCAAGATTATTCTCTGCCCGCAGCTCTCGAGT 730  
 QY 240 SerTrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeuValGluMet 259  
 Db 731 AGCTGGGACTACAGGCGGCCACCAAGCTAGCTAATTTTGTATTTTATTTAGTAGAGATG 790  
 QY 260 -GlyPheThrMetPheAlaArgLeuLeuLeu-IleSerGlyProCysAspLeuProAlaAs 279  
 Db 791 GGGTTTCAACATGTTCCCGCAGGTTGATGTAGATCTCTTGACCTGTGTGATCTGCTGCTC 850  
 QY 279 erAlaSerGlnSerAlaGlyIleThrGlyValSerHisHis-AlaArgLeuIlePheAsn 298  
 Db 851 CGGCTCCCAAAGTCTGGGATTACAGGACGTGACGCCCGCCCGGCTATTTTAAAT 910  
 QY 299 PheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsn 318  
 Db 911 TTTTGTGTTGTTGAAATGGAATCTCACTCTGTATCCAGGCTGGAGTGCATGGCCAAAT 970  
 QY 319 LeuGlySerLeuGlnProLeuProProGlyLeuLysArgPheSerCysLeuSerLeuPro 338

Db 971 CTCGGCTCACTGCAACCTCTGCTCCCGGGCTCAAGCAGATTCTCCTGTCTCAGCCTCCA 1030  
 Qy 339 SerSerTrpAspTyrGlyHisLeuProHisProAlaAsnPheCysIlePheLeuArg 358  
 Db 1031 AGCAGCTGGGATTACGGGACCTGCA-CCACACCCCGCTAAATTTTGTATTTTCATTAGA 1089  
 Qy 359 GlyGlyValSerProTyrLeuSerGlyTyrSerGlnThrProAspLeuArg 375  
 Db 1090 GCGGGGTTTACCATATTTTGACGCTGGTCTCAAACTCTGACCTCAGG 1140

## RESULT 6

US-08-454-557C-49  
 ; Sequence 49, Application US/08454557C  
 ; Patent No. 5830670  
 ; GENERAL INFORMATION:  
 ; APPLICANT: de la Monte, Suzanne  
 ; APPLICANT: Wands, Jack R.  
 ; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection  
 ; TITLE OF INVENTION: of Alzheimer's Disease  
 ; NUMBER OF SEQUENCES: 121  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
 ; STREET: 1100 New York Avenue, Suite 600  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20005-3934  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/454,557C  
 ; FILING DATE: 30-MAY-1995  
 ; CLASSIFICATION: 514  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Ludwig, Steven R.  
 ; REGISTRATION NUMBER: 36,203  
 ; REFERENCE/DOCKET NUMBER: 0609.3840003  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 371-2600  
 ; TELEFAX: (202) 371-2540  
 ; INFORMATION FOR SEQ ID NO: 49:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1381 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: both  
 ; TOPOLOGY: both  
 ; US-08-454-557C-49

Alignment Scores:  
 Pred. No.: 3,25e-135 Length: 1381  
 Score: 1520.00 Matches: 344  
 Percent Similarity: 90.16% Conservative: 4  
 Best Local Similarity: 89.13% Mismatches: 25  
 Query Match: 74.73% Indels: 22  
 DB: 2 Gaps: 3

US-09-380-203-2 (1-375) x US-08-454-557C-49 (1-1381)

Qy 1 MetGluPheSerLeuLeuProHisProAlaAsnPheCysIlePheLeuArg 20  
 Db 14 ATGGAGTTTTTCGCTCTTGTGTCCCGAGCTGGAGTGCATTCAGCTCAGCTCAGC 73  
 Qy 21 AnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40  
 Db 74 AACCTCCGCTCCCGGTTCAAGCAGATTCTCTGTGCTCAGCTCCCGAGTAGCTGGATT 133  
 Qy 41 ThrGlyMetCysThrHisAlaArgLeuLeuLeuTyrPhePheLeuValGluMetGluPhe 60  
 Db 134 ACAGGCGATGTGCAC-CAGCTCGGCTAAATTTTGTATTTTGTATTTTGTAGTAGAGTGGATT 192

Qy 61 --LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspSerProSerValSerA 80  
 Db 193 AACTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATCTCCGCTCGCGC 252  
 Qy 80 laSerGlnSerAlaArgTyrArgThrGly-----HisHisAlaArg-LeuCysLeuAla 97  
 Db 253 CTGCCCCAAGTGTCTGAGATT---ACAGGATGAGCCACCATGCCCCGGCTCTGCTGGCT 309  
 Qy 98 AnPheCysGlyArgAsnArgValSerLeuMet-CysProSerTrpSerProGluLeu 117  
 Db 310 AATTTTGTGGTAGAAACAGGGTTTCACTGATGTTGCCAACAGCTGGTCTCCTGAGCTCAA 369  
 Qy 117 sGlnSerThrCysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProG 137  
 Db 370 GCAGTCCACCTGCCTCAGCTCCCAAGTGTGGGATTACAGGGCTCA-GCGTGCCTGG 428  
 Qy 137 YLeuPheIleLeuPheLeuArgHisArgCysProThrLeuThrGlnAspGluValG 157  
 Db 429 CTTTATTTATTTATTTTATTTTAAAGACACAGGTGTACCACTCTTACCCAGGATGAAGTGA 488  
 Qy 157 nTrpCysAspHisSerSerLeuGlnProSerThrProGluLeuLys-HisProAla 177  
 Db 489 GTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCTCAGATCAAGCAATCTCTGCT 548  
 Qy 177 erAlaSerGlnValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePhe 197  
 Db 549 CAGCTCCCAAGTAGCTGGGACCAAGACATGACCACTACACCTGGTA-ATTTTATTT 607  
 Qy 197 heIlePheAsnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTPA 217  
 Db 608 TTATTTTAAATTTTGTAGACAGAGTCTCAC-TCGTCAACCCAGGCTGGAGTGCAGTGC 666  
 Qy 217 rGAsnLeuGlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerL 237  
 Db 667 GCAATCTTGGCTCACTGCAACTCTGCTCCCGGTTCAAGTTATTTCTCTGCCCGCAGCC 726  
 Qy 237 euLeuSerSerTrpAspTyrArgArgProProArgLeuAlaAsnPheValPheLeu 257  
 Db 727 TCCTGAGTAGCTGGGACTACAGGCGCCACACGCTAGCTAGTAATTTTGTATTTTAG 786  
 Qy 257 alGluMet-GlyPheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeu 276  
 Db 787 TAGAGATGGGTTTCCACATGTTGCCAGGTTGATCTTGATCTCTTGACCTTGTGATCTG 846  
 Qy 277 ProAlaSerAla-SerGlnSerAlaGlyIleThrGly-ValSerHisHisAlaArgLeu 296  
 Db 847 CTTGCTCGGCTACCCAAAGTGTGGATTACAGTCTGACTCCAGCC---CGGCCTA 903  
 Qy 296 lePheAsnPheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnT 316  
 Db 904 TTTTAAATTTTGTGTTGAAATGGAATCTCACTCTGTGTACCAGGTCGAGAGTGCAAT 963  
 Qy 316 rpProAsnLeuGlySerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeu 336  
 Db 964 GGCA-AATCTCGGCTACTCGCAACCTCTGCTCCCGGTC-AAGCGATTCTCTGTCTCA 1021  
 Qy 336 erLeuProSerSerTrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIleP 356  
 Db 1022 GCCTCCCAAGCAGCTGGGATTACGGAC-CTGCA-CCACACCCCGCTAAATTTTGTATTT 1079  
 Qy 356 heIleArgGlyValSerProTyrLeuSerGlyTyr-SerGlnThrProAspLeuArg 375  
 Db 1080 TCATTAGAGCGGGTTTA--CCATATTTGTGAGCTGGGTCTCAAACTCTCAGCTCAGG 1137

## RESULT 7

US-08-340-426D-49  
 ; Sequence 49, Application US/08340426D  
 ; Patent No. 5948634  
 ; GENERAL INFORMATION:  
 ; APPLICANT: de la Monte, Suzanne  
 ; APPLICANT: Wands, Jack R.  
 ; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection  
 ; TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.,  
 STREET: 1100 New York Avenue, Suite 600  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20005-3934

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/340,426D  
 FILING DATE: 14-NOV-1994  
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
 NAME: Ludwig, Steven R.  
 REGISTRATION NUMBER: 36,203  
 REFERENCE/DOCKET NUMBER: 0609,3840002

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 371-2600  
 TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 49:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1381 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: both

US-08-340-426D-49

Alignment Scores:  
 Pred. No.: 3,25e-135 Length: 1381  
 Score: 1520.00 Matches: 344  
 Percent Similarity: 90.16% Conservative: 4  
 Best Local Similarity: 89.12% Mismatches: 25  
 Query Match: 74.73% Indels: 22  
 DB: 2 Gaps: 3

US-09-380-203-2 (1-375) x US-08-340-426D-49 (1-1381)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyValAlaIleSerAlaHisArg 20  
 Db 14 ATGGAGTTTCGCTCTGTTGTCAGGCTGGAGTGAATGGCGCAATCTCAGCTACCGC 73  
 QY 21 AnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40  
 Db 74 AACCTCGGCTCCCGGGTTCAGCGGATCTCTGCTCAGCCTCCCGCAGTAGTGGGATT 133  
 QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60  
 Db 134 ACAGGATGTGAC-CACGCTCGGCTAAATTTGTAATTTTGTAGTAGAGTGGATT 192  
 QY 61 --LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspSerProValSerA 80  
 Db 193 AACTCCATGTGCTCAGGCTGGTCTGAACTCCGACCTCAGATGATCTCCGCTCGGC 252  
 QY 80 laSerGlnSerAlaArgTyrArgThrGly-----HisHisAlaArg-LeuCysLeuAla 97  
 Db 253 CTGCCAAAGTGTAGATT--ACAGGATGAGCCACCATGCGCGCTCTGCTGGCT 309  
 QY 98 AnPheCysGlyArgAsnArgValSerLeuMet-CysProSerTrpSerProGluLeu 117  
 Db 310 AATTTTGTGGTGAAGAGGGTTCACGTGATGTGCGCAAGCTGCTCCTGAGCTCAA 369  
 QY 117 sGlnSerThrCysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProG 137  
 Db 370 GCAGTCCACTGCTCAGCTCCCAAGTGTGGGATTCAGCGCTCA-GCGTGGCTGG 428  
 QY 137 YLeuPheLeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValG 157  
 Db 429 CCTTTTATTTTATTTTATTTTAAAGACACAGGTGTACCACTCTTACCAGGATGAAGTGA 488

QY 157 nTrpCysAspHisSerSerLeuGlnProSerThrProGluIleLys-HisProProAla 177  
 Db 489 GTGGTGTGATCACAGCTACTGAGCTTTCAACTCTGAGATCAAGCAATCTCTCGCCT 548  
 QY 177 erAlaSerGlnValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIleP 197  
 Db 549 CAGCCTCCCAAGTAGCTGGGACCAAGACATGACACCTACACCTGGTA-ATTTTATTT 607  
 QY 197 heIlePheAsnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpA 217  
 Db 608 TTATTTTAAATTTTGTAGACAGAGTCTCAC-TCTGTCCACCGAGGTGGAGTGCAGTGC 666  
 QY 217 iqAsnLeuGlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerL 237  
 Db 667 GCAATCTTGGCTCACTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCCAGCC 726  
 QY 237 euLeuSerSerTrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeu 257  
 Db 727 TCCTGAGTAGCTGGGACTACAGCGCCACACCGCTAGCTAATTTTGTATTTTAG 786  
 QY 257 alGluMet-GlyPheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeu 276  
 Db 787 TAGAGATGGGGTTTCAACATGTTCCAGGTTGATCTTGTATCTTGTGATCTG 846  
 QY 277 ProAlaSerAla-SerGlnSerAlaGlyIleThrGly-ValSerHisHisAlaArgLeu 296  
 Db 847 CTGCTCGGCTCACTCCCAAGTCTGGGATTCAGGTCTGACTCCACGC---CGGCCTA 903  
 QY 296 lePheAsnPheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGln 316  
 Db 904 TTTTAAATTTTGTGTTTGAATGGAATCTCACTCTGTACCAGGTCGGAGTGCAAT 963  
 QY 316 rpProAsnLeuGlySerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeu 336  
 Db 964 GGCA-AATCTCGGCTACTCGCAACCTCTGCTCCCGGGTC-AAGCGATTCTCTGTCTCA 1021  
 QY 336 erLeuProSerSerTrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIleP 356  
 Db 1022 GCCTCCCAAGCAGCTGGGATTCAGGAC-CTGCA-CCACACCCCGTAAATTTTGTATT 1079  
 QY 356 heIleArgGlyGlyValSerProTyrLeuSerGlyTrp-SerGlnThrProAspLeuArg 375  
 Db 1080 TCATTAGAGCGGGTTTA--CCATATTGTGAGGTGGTCTCAAACTCTCACTCAGCTCAG 1137

## RESULT 8

US-08-450-673C-49  
 Sequence 49, Application US/08450673C  
 Patent No. 5948888

GENERAL INFORMATION:  
 APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.

TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/450,673C

FILING DATE: 30-MAY-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

```

; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1381 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-450-673C-49

Alignment Scores:
Pred. No.: 3,25e-135 Length: 1381
Score: 1520.00 Matches: 344
Percent Similarity: 90.16% Conservativity: 4
Best Local Similarity: 89.12% Mismatches: 25
Query Match: 74.73% Indels: 22
DB: 2 Gaps: 3

US-09-380-203-2 (1-375) x US-08-450-673C-49 (1-1381)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
DB 14 ATGGAGTTTCGCTCTTGTGGCCAGGCTGGAGTGCATGGCGCAATCTCAGCTCACGCC 73

QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
DB 74 AACCTCCGCTCCCGGGTCAAGCGATTCTCTGCGCTCAGCTCCCGAGTAGCTGGATT 133

QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60
DB 134 ACAGGCAATGTCAC-CACGCTCGGCTAAATTTGTATTTTATTTTATTTAGTAGAGTAGATT 192

QY 61 --LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspSerProSerValSerA 80
DB 193 AACTCATGTTGGTACGGCTGGTCTCGAATCCCGACCTCAGATGATCTCCGCTCGGC 252

QY 80 laSerGlnSerAlaArgTyrArgThrGly-----HisHisAlaArg-LeuCysLeuAla 97
DB 253 CTGCCCAAAGTGTGAGATT--ACAGGCATGAGCCACCATGCCGCGCTCTGCTGGCT 309

QY 98 AsnPheCysGlyArgAsnArgValSerLeuMet-CysProSerTrpSerProGluLeuIly 117
DB 310 AATTTTGTGTAGAACAGGGTTTCACTGATGTTGCCCAAGCTGGTCTCTCTGAGCTCAA 369

QY 117 sGlnSerThrCysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGl 137
DB 370 GCAGTCCACTGCTCAGCTCCCAAAGTGGGATACAGCGGTCA-GCCGTGCGCTGG 428

QY 137 YLeuPheIleLeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGl 157
DB 429 CCTTTTATTTTATTTTATTTTAAAGACACAGGTGTACCACTTTACCAGGATGAAGTGA 488

QY 157 nTrpCysAspHisSerSerLeuGlnProSerThrProGluIleLys-HisProProAlaS 177
DB 489 GTGGTGTATCATCAGCTCAGCTCACTGAGCTTTCAACTCTGTGATCAAGCAATCTCTGCT 548

QY 177 eAlaSerGlnValAlaGlyThrLysAspMetHisIstYrThrTrpLeuIlePheIleP 197
DB 549 CAGCTTCCCAAGTAGCTGGGACCAAGATGACCATGACCATGACCTGCTGCTGCTGCTGCT 607

QY 197 heilePheAsnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpA 217
DB 608 TTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 666

QY 217 rGAsnLeuGlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerL 237
DB 667 GCAATCTGGCTCAGCTGCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCTCCCGCC 726

QY 237 euLeuSerSerTrpAspTyrArgArgProProArgLeuAlaAsnPheValPheLeuV 257

```

```

DB 727 TCCTGAGTAGCTGGGACTACAGGGCCCAACAGGCTAGCTAAATTTTTTTGTTATTTTAG 786
QY 257 alGluMet-GlyPheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeu 276
DB 787 TAGAGATGGGGTTTCACCATGTTGCCAGGTTGATCTTTGATCTCTTGACCTTTGTGATCTG 846
QY 277 ProAlaSerAla-SerGlnSerAlaGlyIleThrGly-ValSerHisAlaArgLeuI 296
DB 847 CCTGCTCGGCTACCCAAAGTGTGGGATTACAGTCTGAGCTCCACGC---CGGCCTA 903
QY 296 lePheAsnPheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTr 316
DB 904 TTTTAAATTTTGTGTTGTAATGGAATCTCACTCTGTACCAGGTGGAGTGCAT 963
QY 316 rpProAsnLeuGlySerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuS 336
DB 964 GGCA-AATCTCGGCTACTCGCAACCTCTGCTCCCGGTC-AAGCGATTCTCTGCTCTCA 1021
QY 336 erLeuProSerSerTrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIleP 356
DB 1022 GCCTCCCAAGCAGCTGGGATTACGGGAC-CTGCA-CCACACCCCTAATTTTGTATTT 1079
QY 356 heileArgGlyGlyValSerProTyrLeuSerGlyTrp-SerGlnThrProAspLeuArg 375
DB 1080 TCATTAGAGCGGGTTTA--CCATATTTGTAGGCTGGGTCTCAAACTCTGACCTCAGG 1137

RESULT 9
PCT-US95-17111A-49
; Sequence 49, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1381 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
PCT-US95-17111A-49

Alignment Scores:
Pred. No.: 3,25e-135 Length: 1381
Score: 1520.00 Matches: 344
Percent Similarity: 90.16% Conservativity: 4

```

Best Local Similarity: 89.12% Mismatches: 25  
Query Match: 74.73% Indels: 22  
DB: 5 Gaps: 3

US-09-380-203-2 (1-375) x PCT-US95-17111A-49 (1-1381)

```
QY 1 MetGluPheSerLeuLeuProArgLeuGluCysAsnGlyValAlaIleSerAlaHisArg 20
DB 14 ATGGAGTTTTCGCTCTGTGTCAGGCTGGAGTGAATGGCGCAATCTCAGCTACCGC 73
QY 21 AnnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
DB 74 AACCTCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCCAGCTCCAGTAGCTGGATT 133
QY 41 ThrGlyMetCysThrHisAlaArgLeuLeuLeuTyPhePheLeuValGluMetGluPhe 60
DB 134 ACAGGATGTGCAC-CACGCTGGCTAAATTTGTATTTTTTTTAGTAGAGTAGGATT 192
QY 61 --LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspSerProValSerA 80
DB 193 AACTTCATGTGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCTCCGCTCGGC 252
QY 80 laSerGlnSerAlaArgTyArgThrGly-----HisHisAlaArg-LeuCysLeuAla 97
DB 253 CTGCCCAAAAGTGTGAGATT---ACAGGCATGAGCCACCATGCCGCTCTGCTGGCT 309
QY 98 AsnPheCysGlyArgAsnArgValSerLeuMet-CysProSerTrpSerProGluLeuIly 117
DB 310 AATTTTGTGTAGAAACAGGGTTTCACGTGTTGCCCAAGCTGCTCTCCAGCTCMA 369
QY 117 sGlnSerThrCysLeuSerLeuProLysCysTrpAspTyArgArgAlaAlaValProG 137
DB 370 CGAGTCCACCTCGCTCAGCCTCCCAAAGTGTGGGATTACAGCGCTCA-GCCGTGCTG 428
QY 137 yLeuPheLeuLeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValG 157
DB 429 CCTTTTATTTATTTTATTTTATTTTAAAGACACAGGTGTACCACTCTTACCCAGGATG 488
QY 157 nTrpCysAspHisSerSerLeuGlnProSerThrProGluLeuLys-HisProAlaAs 177
DB 489 GTGGTGTATCAAGCTCACTGAGCCTTCAACTCTGAGATCAAGCAATCTCTCTGCT 548
QY 177 exAlaSerGlnValAlaGlyThrLysAspMetHisHisTyThrTrpLeuIlePheIleP 197
DB 549 CAGCCTCCCAAGTAGTGGGACCAAGACATGACCACTACACCTGCTA-ATTTTATTT 607
QY 197 heIlePheAsnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpA 217
DB 608 TTATTTTAAATTTTGTAGACAGAGTCTCAC-TCTGTCAACCCAGGCTGGAGTGCAGTGC 666
QY 217 tGAsnLeuGlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerL 237
DB 667 GCATCTTGCTCACTGCAACCTCTGCTCCCGGTTCAAGTTATCTCTCCGCCAGCC 726
QY 237 euLeuSerSerTrpAspTyArgArgProProArgLeuAlaAsnPhePheValPheLeuV 257
DB 727 TCCTGAGTAGCTGGACTACAGCGCCCAACCCAGCTAGCTAATTTTGTATTTTAG 786
QY 257 alGluMet-GlyPheThrMetPheAlaArgLeuLeuLeuIleSerGlyProCysAspLeu 276
DB 787 TAGAGATGGGGTTTCAACATGTTCCGAGTTGATCTTGTATCTTGTACCTTGTGATCTG 846
QY 277 ProAlaSerAla-SerGlnSerAlaGlyIleThrGly-ValSerHisHisAlaArgLeuI 296
DB 847 CTGCTCGCTCGCTTACCAAGTGTGGGATTACAGGTGCTGACTCCAGCC---CGGCCTA 903
QY 296 lePheAsnPheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnT 316
DB 904 TTTTAAATTTTGTGTTGTAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAT 963
QY 316 rpProAsnLeuGlySerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuS 336
DB 964 GCAC-AATCTCGGCTACTCGCAACCTCTGCTCCCGGGTC-AAGCGATTCTCTCTCTCA 1021
```

```
QY 336 erLeuProSerSerTrpAspTyArgGlyHisLeuProProHisProAlaAsnPheCysIleP 356
DB 1022 GCCTCCCAAGCAGCTGGGATTACGGGAC-CTGCA-CCACACCCCGCTAAATTTTGTATT 1079
QY 356 heIleArgGlyGlyValSerProTyArgLeuSerGlyTrp-SerGlnThrProAspLeuArg 375
DB 1080 TCATTAGAGCGGGTTTA--CCATATTCTCAGGCTGGGTCTCAAACTCCTGACCTCAGG 1137
```

## RESULT 10

```
US-09-949-016-15370
; Sequence 15370, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15370
; LENGTH: 9779
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15370
```

## Alignment Scores:

```
Pred. No.: 4,428-76 Length: 9779
Score: 913.00 Matches: 239
Percent Similarity: 59.76% Conservative: 15
Best Local Similarity: 56.24% Mismatches: 112
Query Match: 44.89% Indels: 62
DB: 4 Gaps: 10
```

US-09-380-203-2 (1-375) x US-09-949-016-15370 (1-9779)

```
QY 4 SerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArgAsnLeuArg 23
DB 3685 TCACCTCTTGTGCTCCAGGCTGGAGTGCATGTGCAATCTCGGCTCACCGCAATCTCGC 3744
QY 24 LeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIleThrGlyMet 43
DB 3745 CTCCTGGGTTCAACAAATTTCTCTGCTCAGCTCCCGAGTAGCTGGGATTATAGGCATA 3804
QY 44 CysThrHisAlaArgLeuLeuLeuTyPhePheLeuValGluMetGluPheLeuHisVal 63
DB 3805 CGCCACCATGCTGGCTAATTTGTAT---TTTTA-GTAGAGACGGGGTTCTCCATGTT 3860
QY 64 GlyGlnAlaGlyLeuGluLeuProThrSerAspSerProSerValSerAlaSerGlnSer 83
DB 3861 GGTCAAGCTGATCTGAACCTCCCACTCCAGCTGATCCACCTGCTCAGCTCCCAAGT 3920
QY 84 AlaArgTyArgThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCysGlyArgAsn 103
DB 3921 CTTGGAATTCAGCATGAGCCACTGCG-----GCCTGGCTTAATTTTGTATTTTAAAT 3974
QY 104 ArgValSerLeuMetCysProSer---TrpSer-----ProGluLeuLysGlnSerThr 120
DB 3975 AGGGTT---TGTGTCATGTTGGTCACTGCTGCTCAAAACCCCTCACATCAACTGATCTGCC 4031
QY 121 CysLeuSerLeuProLysCysTyPheAspTyArgArgAlaAlaValProGly-----Leu 138
DB 4032 TGCCTCAGCTTCCAAAGTGTGGGATTACAGGGGTGA-GCCATCGCACATAACTTCCTT 4090
QY 139 PheIleLeuPhePheLeuArgHisArgCysPro---ThrLeu-ThrGlnAspGluValG 157
```

[illegible]

RESULT 11

```

US-09-949-016-14370/c
; Sequence 14370, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

```

```

QY 176 AlaSerAlaSerGlnValAlaGlyThrLysAspMetHisHisThrTrpLeuIlePhe 195
Db 22741 GCTTAGCCTCTGAGTAGCTGGATTACAGGCACGCCACACCGCCAGGCTAATTTT 22682
QY 196 IlePheIle----- 198
Db 22681 GTATTTTGTAGTAGACGGGTTTCCACCAATTTTGGCCAGGCTGGTCTTGAATTCCTGACT 22622
QY 198 ----- 198
Db 22621 TCGTATCGCCTGCTCAGCCTCCCAAGAGTCTGGGATTACAGCGGTGAGCCACACAC 22562
QY 199 -----PheAsnPheLeuArgGlnSerLeuAenSerValThrGlnAlaGlyValGln 215
Db 22561 CTGGCCTATTATTTTGTAGATGAGTCTTGC-TCTGTTGCCAGGCTGAGTGCAG 22503
QY 216 TrpArgAsnLeuGlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysPro 235
Db 22502 TGGTGCCATCTTGGCTCACTGCAACCTCCGCCTCTGGGTTTCAAGTGATTCTCTGCCTC 22443
QY 236 SerLeuLeuSerSerTrpAspTyrArgArg---ProProArgLeuAlaAsn----- 251
Db 22442 ACGTCCGAGTAGCTGGATTACAGCGCCAGCCACCATGCGCGCTAATTTTGTATT 22383
QY 252 ----PhePheValPheLeuValGluMetGlyPheThrMetPheAla---ArgLeu----- 267
Db 22382 TCTTTTTTTTTTTTTTTTTTGTAGATGGAGTCTCACTCTTGTGCCCGAGGCTGAGTGC 22323
QY 268 -----IleLeuIleSerGlyProCysAspLeuProAla 278
Db 22322 AACAGTGGCATCTTGGCTCACTGCAACCTCTGCCTCCCGGTTTCAAGCGATTCTCTGCC 22263
QY 279 SerAlaSerGlnSerAlaGlyIleThrGlyValSerHisHisAlaArgLeuIlePhe-As 298
Db 22262 TCAGCTCCGAGTAGCTGGGATTACAGCGCAAGCCACCATGCGCGCTAATTTTGTGA 22203
QY 298 nPheCysLeu-----PheGluMetGluSerHisSer-ValThrGlnAlaGlyValG 315
Db 22202 TTTCTTTCTTTTTTTTTTTTGTAGATGGAGTCTCACTCTTGTCACTAGGCTGAGTGC 22143
QY 315 InTrpProAsnLeuGlySerLeuGlnProLeuProGlyLeuLysArgPheSerCysL 335
Db 22142 AACGGTGGCATCTTGGCTCACTGCAACCTCTGCCTCCCGGTTTCAAGCGATTCTCTGCC 22083
QY 335 euSerLeuProSerSerTrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysI 355
Db 22082 TCAGCCTCTTGTAGTACCTGGGATTATAGATGCTGCCACCATGCGCGCTAATTTTGTGA 22023
QY 355 IlePheIleArgGlyGlyValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuA 375
Db 22022 TTTGTAGTAGACAGCGGTTTCCACCATATTGAACAGGCTGGTCTCGAACTCTTGACCTCA 21963
QY 375 rg 375
Db 21962 GG 21961

RESULT 12
US-09-949-016-12315/c
; Sequence 12315, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

```

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12315
; LENGTH: 50518
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12315

Alignment Scores:
Pred. No.: 2,418-72 Length: 50518
Score: 885.00 Matches: 236
Percent Similarity: 51.86% Conservative: 29
Best Local Similarity: 46.18% Mismatches: 101
Query Match: 43.51% Indels: 147
DB: 4 Gaps: 12

US-09-380-203-2 (1-375) x US-09-949-016-12315 (1-50518)
QY 4 SerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArgAsnLeuArg 23
Db 39610 TCACCCCTTTGTTGCCAGGCTGAGTGCAATGGCGTGGTCTCGGCTCATGGCAACATCTGC 39551
QY 24 LeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIleThrGlyMet 43
Db 39550 CTCCCGGTTTCAAGCCATTCTCTGCTCAACCTCCCAAGTAGCTGGGATTACAGGTGTG 39491
QY 44 CysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPheLeuHisVal 63
Db 39490 CGCCTATGATGCCAGCTAATTTTATATATTTTGTAGATGAGGTTCACCATGTT 39434
QY 64 GlyGlnAlaGlyLeuGluLeuProThrSerAspProSerValSerAlaSerGlnSer 83
Db 39433 GGCCAGGCTTCTGCAGAACTCTGACCTGAGGTGATCCACCCACCTCAGCCTCCCAAGT 39374
QY 84 AlaArgTyrArgThr---GlyHisHisAlaArg-Leu----- 94
Db 39373 GCTGGGATTACAAGCATGAGCCACCGCGCGCTATATTTATTTATTTTATTTTATTT 39314
QY 94 ----- 94
Db 39313 ACTTATTTATTTGACTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 39254
QY 94 ----- 94
Db 39253 GTCTTGCTCTCTCACCCAGGCTGAAGTCAATGGCATGCTCTGCTCCTCACTGCAACCCCC 39194
QY 94 ----- 94
Db 39193 ACTTCCCAAGTTCAAGCAATTTCTCTGCTCAGCCTCCCGAGTAGCTGGGACTACAGGTG 39134
QY 95 -----CysLeuAlaAsnPheCys-----GlyArgAsnArgValSerLeuMetCys 109
Db 39133 TGTGCCACCATGCTGGCTAATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGT 39082
QY 109 sProSerTrp-----SerProGluLeuLysGlnSerThrCysLeuSerIle 124
Db 39081 -CCATCTGGCCAGGCTAATCTCGAATCTCTGAGCTCAGGTGATCGGCCCTCGGCT 39023
QY 124 uProLysCysTrpAspTyrArgArgAlaVal---ProGlyLeuPheIleLeuPhePhe 143
Db 39022 CCCAAATGCTGAGATTACAGCGGTGAGCCACACACCGCTCTTTTGT---TTTT 38966
QY 143 eLeu-ArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAspHisSerS 163
Db 38965 TTTGAGATGGGATCTCACTCTTTCACCCAGCCAGAGTACAGTGGAGAAATCATAGCT 38906
QY 163 erLeuGlnProSerThrProGluIleLys-HisProProAlaSerAlaSerGlnValAla 182
Db 38905 CACTGCAGCATCGAACTCTGAGCTCAAGGATCTCTCCAGCTCAGCCACCTAGTAGCT 38846
QY 183 GlyThrLysAspMetHisHisThrTrpLeuIlePheIlePheLeuPheLeu 202
Db 38845 GGGATCAAAAGGTGTGCACCACTGCCAGGCTAATTTGTTTAA-TTTTTTTTTTTTTT 38787

```



```
QY 203 ArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGlySerLeu 222
Db 38786 AGATGGAGTCTT---GCTGTACCCAGCGCTGGAGTGCAGTGGTCAATCTCAGCTCACTG 38730
QY 223 GlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSerTrpAsp 242
Db 38729 CAACTCTGCCCTCCAGGGTTCAAGCAGTTCTCTGCTCAGCCTCTTGAGTAGTAGGAC 38670
QY 243 TyrArg---ArgProProArgLeuAlaAsnPhePheValPheLeuValGluMetGlyPhe 261
Db 38669 TACAAGTGCCCAACACACACCTGGCTAATTTTGTATTTTAGTAG-AGACAGGGTTTC 38611
QY 262 ThrMetPheAlaArgLeuLeuLeuSerGlyProCysAspLeuProAlaSerAlaSer 281
Db 38610 ACCATGTTGGTCCAGGCTGCTCGAACTCTCGAACCTTGATCTGCTCCTCCTCAGCCTCC 38551
QY 282 GlnSerAlaGlyIleThrGlyValSerHisHisAla----- 293
Db 38550 CAAAGTCTGGGATTACAGGTGTGAGCCACCGTGCCTGGCTGTTTAAATTTTTTAGAGA 38491
QY 294 -----ArgLeuIlePheAsn----- 298
Db 38490 CGGGGTCTTGCTATGTGGCCAGCGCTGGTCTCGAACTCTCGGCTCAGGCAATCCTTCCA 38431
QY 298 ----- 298
Db 38430 CCTCAGCTATCTGAGAGTAAGTGGGATTATAGGGAAGAGCCATGGCAGCTGGTGGCTT 38371
QY 299 -----PheCysLeuPheGluMetGluSer 306
Db 38370 GTTACATTTTAACTCCTATTGCTCTCTACTCTCTTTTCTTTTCTCGAGACAGTCT 38311
QY 307 HisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGlySerLeuGlnProLeuPro 326
Db 38310 TGTCTCGTCCCCAGCGCTGGAGTACAGTGTGCGATCTCAGCTCACTGCAACCTCGGCT 38251
QY 327 ProGlyLeuLysArgPheSerCysLeuSerLeuProSerSerTrpAspTyrGlyHisLeu 346
Db 38250 CTGGGTTCACGGGTTCCTCGCTCAGCTCCGAGTAGTGAATAATTACAGACATTC 38191
QY 347 ProProHisProAlaAsnPheCysIlePheIleArgGlyValSerProTyrLeuSer 366
Db 38190 ACACATGCCAGCTAATTTTGTATTTTAGTAGAGACAGGGTTTCGCCATGTAGCCA 38131
QY 367 GlyTrpSerGlnThrProAspLeuArg 375
Db 38130 GACTAGTCTCAAACTCCTCAGCTCAGG 38104
```

## RESULT 13

```
US-09-949-016-13718
; Sequence 13718, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13718
; LENGTH: 14411
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13718
```

```
Alignment Scores:
Pred. No.: 7,13e-72 Length: 14411
Score: 871.50 Matches: 250
Percent Similarity: 45.62% Conservative: 21
Best Local Similarity: 42.09% Mismatches: 99
Query Match: 42.85% Indels: 231
DB: 4 Gaps: 9

US-09-380-203-2 (1-375) x US-09-949-016-13718 (1-14411)
QY 1 MetGluPheSerLeuLeuProArgGluCysAsnGlyAlaIleSerAlaHisArg 20
Db 5680 ATGGAGTTTG-CTCTGTGCCAGGTGGAGTGAATGGTGCATCTTGGCTCACTGC 5738
QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40
Db 5739 AACCTCCGCTCCAGGTTCAAGCAATTTCTCTGCTCAGTCCCCCAAGTAGCTGGATT 5798
QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuPhePheLeuValGluMetGluPhe 60
Db 5799 ACAGCAGCTGCGCACACACCCAGCTAATTT--TTGTATTTTATAGTAGACGGGTTT 5856
QY 61 LeuHisValGlyGlnAlaGlyLeuLeuProThrSerAspProSerValSerAla 80
Db 5857 CTCCATGTGTGTCAGGTGGTCTCAAACTCTCAGCTCAGGTGATCCACCGCCTCGGCC 5916
QY 81 SerGlnSerAla-----ArgTyrArgThr----- 88
Db 5917 TCCCAAAATGTGGGATTACAGGCATGAGCCAGCGCTGGCCTCCCCACCTCTTAT 5976
QY 88 ----- 88
Db 5977 GTCCTGTCTGGAACACAGCACGCCGCTCTCCCCAGTAGAGAAGTGACGCCACTAATCGG 6036
QY 89 GlyHisHisAlaArgLeuCysLeuAlaAsnPheCysGlyArgAsnArgVal---SerLeu 107
Db 6037 GGTATACATCTTTGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTT 6096
QY 108 MetCysProSerTrpSer----- 113
Db 6097 CGTTGCCAGCGTGGAGTGCAATGTCAGCATCTCGGCTCACTGCAACCTCTTCTGCCTGC 6156
QY 113 ----- 113
Db 6157 TGGGTTCAAGCAATTTCTCTGTCTCAGCCTCACTAGTAGTGGGATTACAGSCACCGCC 6216
QY 113 ----- 113
Db 6217 ACCATGCCAGCTAATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTT 6276
QY 114 -----ProGluLeuLysGlnSerThrCysLeuSerLeuPro 125
Db 6277 TGTGATCAGCGTGTATCTTGAACCTCTGACCTCAGGTGATCCACCCACCTCAGCCTCCA 6336
QY 126 LysCysTrpAspTyrArgArg--AlaAlaValProGlyLeu----- 138
Db 6337 AAGTCTGGAATACAGGCGTGAGCCACTGTGCCGCGCTCACATCTTTGTTTGTATTTTACCA 6396
QY 138 ----- 138
Db 6397 CCATGTTTATTTTATTTTATTTATATATATATATATATATATATATATATATATATATAT 6456
QY 138 ----- 138
Db 6457 TGTTCCTTGTGCCCAAGTTGGAGTGAATGGCAGCATCTCAGCTCATTGCAACCTCTGC 6516
QY 138 ----- 138
Db 6517 CTCCCGGTTCAAGTGATTCTCTCGCTCAGCTCCTGAGTAGTGGGAAATACAGCGCT 6576
QY 139 -----PheIleLeuPhePheLeuArgHisArgCysProThrL 151
Db 139 -----PheIleLeuPhePheLeuArgHisArgCysProThrL 151
```

```
Db 6577 CGCACACACACCAGCTAAATTTTTTTTTTTTTTTTTTTTTTTTGTGACAGGAGTCTCGCTTT-G 6635
Qy 151 euThrGlnAspGluValGlnTrpCysAspHisSerSerLeuGlnProSerThrProGluI 171
Db 6636 TCGCCAGGTGGAGTGCAGTGTGCAATCTGACTCACTGCAACCTCGCCTCCAGCT 6695
Qy 171 leuYHisProProAlaSerAlaSerGlnValAlaGlyThrLysAspMetHisHisTyr 191
Db 6696 TCAAGCACTTCTCTGCTCAGCTGCGAGTAGCTGGATTACAGTGGCTGCCACCATG 6755
Qy 191 hrTrpLeuLeuPheLeuPhe----- 197
Db 6756 CTGGCTAAATTTTTTT-TTTTTGTATTTTGTAGTACAGAGGGTTTTTACCATCTGGCCA 6814
Qy 197 ----- 197
Db 6815 GGCTGCTCTTAACTCTGACCTCGTGTGATCCACCCCTCAGCTCCCAAGTGTGAGA 6874
Qy 198 -----IlePheAsnPheLeuArgGlnSerLeuA 207
Db 6875 TTACAGGTGTAAGCCACCATGCTGCGCAATTTTATTTATTTTGTAGACAGAGTCTCG 6934
Qy 207 snSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGlySerLeuGlnProLeuProp 227
Db 6935 C-TCTGTGCCAGGTGGAGTGCAGTGTGATACAACTTGCTCACTGCAACCTCGCCTC 6993
Qy 227 roGlyPheLysLeuPheSerCysProSerLeuLeuSerSerTrpAspTyrArgAtg---P 246
Db 6994 CCGGGTTAAGCAATTTCTCTGCTCAGCTTCTGAGTAGCTGGAGTACAGCGCGTGC 7053
Qy 246 roProArgLeuAlaAsnPhePheValPheLeuValGluMetGly-PheThrMetPheAla 265
Db 7054 CACCATGCCCGGCTAA-TTTTTGTGTTTGTAGTACAGACAGGATTTTACCATTGTGGCC 7112
Qy 266 ArgLeuLeuLeuLeuSerGlyProCysAspLeuProAlaSerAlaSerGlnSerAlaGly 285
Db 7113 AGGCTGGTCTCGATCTCTGACCTCATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7172
Qy 286 IleThrGlyValSerHisAlaArgLeuLeuPheAsn-----PheCysLeuPheGlu 303
Db 7173 ATTACAGGGTGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7231
Qy 304 MetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGlySerLeuGln 323
Db 7232 ATGGAGTCTCGCTCTGTGCGCCAGGTGGAGTGCAGTGGCGTGTGCTGCTGCTGCTGCTG 7291
Qy 324 ProLeuProProGlyLeuLysArgPheSerCysLeuSerLeuProSerSerTrpAspTyr 343
Db 7292 GCTCCGCTCTCTGGGTTCACACCATTTCTCTGCTCAGCTCCCGAGTAGCTGGGACTAC 7351
Qy 344 GlyHisLeuProProHisProAlaAsn-PheCysIlePheLeuArgGlyGlyValSerPr 363
Db 7352 AGGTGCCCGCCAGCGCCAGCTAAATTTTGTATTTTGTAGTACAGTGGGTTTACC 7411
Qy 363 oTyrLeuSerGlyTrpSerGlnThrProAspLeu 374
Db 7412 G-----TGCTCTATCTCTCTGACCTC 7433
RESULT 14
US-09-949-016-12462/c
; Sequence 12462, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```

```
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12462
; LENGTH: 56616
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(56616)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12462

Alignment Scores:
Pred. No.: 1,078-70 Length: 56616
Score: 868.50 Matches: 231
Percent Similarity: 57.04% Conservative: 16
Best Local Similarity: 53.35% Mismatches: 94
Query Match: 42.70% Indels: 96
DB: 4 Gaps: 12

US-09-380-203-2 (1-375) x US-09-949-016-12462 (1-56616)
Qy 5 LeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArgAsnLeuArgLeu 24
Db 45864 CTCCTTTGTCGCCAGACTGGAGTGCACGGCGTGATCTCAGCTCATTTGCAACCTCTGCTTC 45805
Qy 25 ProGlySerSerAspSerProAlaSerAlaSerProAlaAlaGlyIleThrGlyMetCys 44
Db 45804 CCAGGTTTCAAGTGATTTCTCTGACTCAGCTCCCAAGTAGCTGGGATTAAGGCGGCACAC 45745
Qy 45 ThrHisAlaArgLeuLeuLeuTyrPhePheLeuValGluMetGluPheLeuHisValGly 64
Db 45744 CACCACACCACTA-----TTTTTAGTAGAGATGGGTTTTCACGCTGTGTG 45697
Qy 65 GlnAlaGlyLeuGluLeuProThrSerAspProSerValSerAlaSerGlnSerAla 84
Db 45696 CAGGCTGGTCTGAACTCTGACCTCAGCTGATCGCTCGCTCGCTCCCAAGTGTCT 45637
Qy 85 ArgTyrArgThrGly-----HisHisAlaArgLeuCysLeuAlaAsnPheCysGlyArg 102
Db 45636 GGGATT---ACAGGATGAGCCACCAACGCT----- 45610
Qy 103 AsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThrCysLeu 122
Db 45609 -----TGG-----CCCCAGTAGAAGCGCTACGTGTACT 45580
Qy 123 -----SerLeuProLysCysTrpAspTyrArgArg----- 132
Db 45579 CCCCATTTCATCATTAAGGATATTACAAAGAGTTGCAATTTTACAGACTAATATTTTAT 45520
Qy 133 -----AlaAlaValProGlyLeuPheLeuPhePheLeuArgHisArgCysPro 149
Db 45519 CACCTCATCAAGGCCATTCTCT-----TTTTTTTTTTTGTAGATG-GAGTATCAC 45473
Qy 150 ThrLeuThrGlnAspGluValGlnTrpCysAspHisSerSerLeuGlnProSerThrPro 169
Db 45472 TCTGTCAACCGAGCTGGAGTGCAGTGGTGCATCTCGGCTCACTGCAAACTCTGCTCTCC 45413
Qy 170 GluIle-LysHisProProAlaSerAlaSerGlnValAlaGlyThrLysAspMetHisH 189
Db 45412 AGTTTCACGCCATTCTCTGCTCAGCCCCCGAGTAGCTGGAGTACAGGTGCCACCA 45353
Qy 189 sTyrThrTrpLeuLeuPheLeuPheLeuPheAsn----- 200
Db 45352 CCACGCCCGGCTAAATTTTTTGT---ATTTTAGTAGACAGCGGGTTTTCACCATGTAGCC 45296
Qy 200 ----- 200
Db 45295 AGGATGGTCTCGATCTCTCTGACCTCATGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 45236
Qy 201 -----PheLeuArgGlnSerLeuAsnSerVa 209
```

```

Db      45235 ATGACAGGATGAGCCACCGTGGCCAGCCTTTTCTTTTGTAGATGAAGTCTCGCTTTGT 45176
QY      209 lThrGlnAlaGlyValGlnTrpArgAsnLeuGlySerLeuGlnProLeuProGlyPh 229
Db      45175 CGCCAGGCTGGAGTACAATGGCACAATCTCGGCTCACTGCAATCTCCGCCCTCTGGGT 45116
QY      229 eLysLeuPheSerCysProSerLeuLeuSerSerTrpAspTyrArgPro---ProAr 248
Db      45115 CRAAGCAATTCTCTGCTCAGCCTCCCAAGTAGTGGGATTACAGGCACCCACTAGCAG 45056
QY      248 gLeuAlaAsnPheValPheLeuValGluMet-GlyPheThrMetPheAlaArgLeu 268
Db      45055 CCAGCTAATT---TTTGTATTTTAGATAGATGGGGTTTACCATTGTTGGCCAGGCTGG 44998
QY      268 leLeuIleSerGlyPro--CysAspLeuProAlaSerAlaSerGlnSerAlaGlyIle 287
Db      44997 TCTCGAATCTCTGACCTCAGGCGATCTGCCACCTCGGCTCCCAAAGTACTGAGATTAC 44938
QY      287 rGlyValSerHisHisAlaArgLeuLe---PheAsnPheCysLeuPheGluMetGl 305
Db      44937 AGCGGTGAGCCACTGCGCCTC-ACCATCAAGGCCATTCTTTTGTGTTT-TTTGAGATGGA 44880
QY      305 uSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGlySerLeuGlnProLe 325
Db      44879 GTCTCACTCTGTGGCCAGCTGGAGTGCAGTGGCAGATCTTGGCTCACTGCAACCTCT 44820
QY      325 uProGlyLeuLysArgPheSerCysLeuSerLeuProSerSerTrpAspTyrGlyHi 345
Db      44819 GCCTCTGGGTTCAAGTGATTTTCTGCTCAGCCTCCCGAGTAGCTGGGATTACAGGCA 44760
QY      345 sLeuProProHisProAlaAsnPheCysIlePheIleArgGlyValSerProTyrIle 365
Db      44759 CCTGCCACCACATCCAGCTGCTTTTGTGTTTGTAGATAGATGGGGTTTCCACCATGTTG 44700
QY      365 uSerGlyTyrSerGlnThrProAspLeuArg 375
Db      44699 ACCAGGATGCTCGAATCTCTGACCGCAGG 44669

RESULT 15
US-09-949-016-17085/c
; Sequence 17085, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 17085
; LENGTH: 56616
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(56616)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17085

```

Alignment Scores:  
 Pred. No.: 1.07e-70 Length: 56616  
 Score: 868.50 Matches: 231  
 Percent Similarity: 57.04% Conservative: 16  
 Best Local Similarity: 53.35% Mismatches: 94

```

Query Match: 42.70% Indels: 96
DB: 4 Gaps: 12

US-09-380-203-2 (1-375) x US-09-949-016-17085 (1-56616)

QY      5 LeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArgAsnLeuArgLeu 24
Db      45864 CTCCTGTTGCCAGACTGAGTGCACCGCGTGATCTCAGCTCATTGCAACCTCTGCTTC 45805
QY      25 ProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIleThrGlyMetCys 44
Db      45804 CCAGGTTCAAGTGATCTCTGACTCAGCTCCCAAGTAGTGGGATTAAAGCGCACAC 45745
QY      45 ThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPheLeuHisValGly 64
Db      45744 CACCACACCCAGCTA-----TTTTTAGTAGATGGGGTTTTCACCTGCGCTTGT 45697
QY      65 GlnAlaGlyLeuGluLeuProThrSerAspProSerValSerAlaSerGlnSerAla 84
Db      45696 CAGGCTGGTCTCGAACTCTGACCTCAGGTGATCCGCTCGCCTGGCCTCCCAAGTGTCT 45637
QY      85 ArgTyrArgThrGly-----HisHisAlaArgLeuCysLeuAlaAsnPheCysGlyArg 102
Db      45636 GGGATT---ACAGGCATGAGCCACCCAGCC----- 45610
QY      103 AsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThrCysLeu 122
Db      45609 -----TCG-----CCCCAGTAAAGAGCGCTACGTGTACT 45580
QY      123 -----SerLeuProLysCysTrpAspTyrArgArg----- 132
Db      45579 CCCCATTTCATCATTAAGGATATTACAAGAGTTGCCAATTTACAGACTAATAATTTTATC 45520
QY      133 -----AlaAlaValProGlyLeuPheIleLeuPheLeuArgHisArgCysPro 149
Db      45519 CACCTCATCAAGGCCATTCT-----TTTTTTTTTTCAGATG-GAGTATCAC 45473
QY      150 ThrLeuThrGlnAspGluValGlnTrpCysAspHisSerSerLeuGlnProSerThrPro 169
Db      45472 TCTGTCAACCCAGGCTGGAGTGCAGTGGTGCAGATCTCGGCTCACTGCAAACTCTGCTCCC 45413
QY      170 GluIle-LysHisProProAlaSerAlaSerGlnValAlaGlyThrLysAspMetHisHi 189
Db      45412 AGGTTACGCCCATTTCTTCTGCTCAGCCCCCGAGTAGCTGGGACTACAGGTGCCACCA 45353
QY      189 sTyrThrTrpLeuIlePheIlePheAsn----- 200
Db      45352 CCACGCCCGGCTAAATTTTGT---ATTTTAGTAGAGAGCGGGTTTCCACCATGTTAGCC 45296
QY      200 ----- 200
Db      45295 AGGATGGTCTCGATCTCTGACCTCATGATCTGCTGCTCGACTCCGACCTCCAAAGTCTAGG 45236
QY      201 -----PheLeuArgGlnSerLeuAsnSerVa 209
Db      45235 ATGACAGGATGAGCCACCGTGGCCAGCCTTTTCTTTTGTAGATGAAGTCTCGCTTGT 45176
QY      209 lThrGlnAlaGlyValGlnTrpArgAsnLeuGlySerLeuGlnProLeuProGlyPh 229
Db      45175 CGCCCGAGGCTGGAGTACAATGGCACAATCTCGGCTCACTGCAATCTCCGCCCTCTGGGT 45116
QY      229 eLysLeuPheSerCysProSerLeuLeuSerSerTrpAspTyrArgPro---ProAr 248
Db      45115 CRAAGCAATTCTCTGCTCAGCCTCCCAAGTAGTGGGATTACAGGCACCCACTAGCAG 45056
QY      248 gLeuAlaAsnPheValPheLeuValGluMet-GlyPheThrMetPheAlaArgLeu 268
Db      45055 CCAGCTAATT---TTTGTATTTTAGTAGATGGGGTTTACCATTGTTGGCCAGGCTGG 44998
QY      268 leLeuIleSerGlyPro--CysAspLeuProAlaSerAlaSerGlnSerAlaGlyIle 287
Db      44997 TCTCGAATCTCTGACCTCAGGCGATCTGCCACCTCGGCTCCCAAAGTACTGAGATTAC 44938

```

```
QY 287 rGlyValSerHisHisAlaArgLeuIle-----PheAsnPheCysLeuPheGluMetGI 305
Db 44937 AGGCGTGAGCCACTGGCCCTC-ACCATCAAGGCCATCTTTTGTGTTT-TTTGAGATGGA 44880
QY 305 uSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGlySerLeuGlnProLe 325
Db 44879 GTCTCACTCTGTGGCCAGGCTGGAGTGCAGTGGCAGAATCTTGGGCTCACTGCAACCTCT 44820
QY 325 uProProGlyLeuLeuArgPheSerCysLeuSerLeuProSerSerTrpAspTyrGlyHi 345
Db 44819 GCCTCTGGGTTCAAGTGATTTTCCTGCTTCAGCCTCCCGAGTAGCTGGGATTTACAGGCA 44760
QY 345 sLeuProProHisProAlaAsnPheCysIlePheIleArgGlyGlyValSerProTyrLe 365
Db 44759 CCTGCCACCACATCCAGCTGCTTTTGTGTTTGTAGTAGAGATGGGGTTTCACCATGTTG 44700
QY 365 uSerGlyTrpSerGlnThrProAspLeuArg 375
Db 44699 ACCAGGATGCTCTCGAACTCTGACCGCAGG 44669
```

Search completed: September 16, 2005, 01:30:19  
Job time : 325 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 15, 2005, 18:46:18 ; Search time 670 Seconds  
(without alignments)  
3313.288 Million cell updates/sec

Title: US-09-380-203-2

Perfect score: 2034

Sequence: 1 MFESLLPLRLECNCAISRAH.....FIRGGVSPYLQSGWQTDLR 375

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q/cn2\_1/USFTO.spool/US09380203/runat\_13092005\_170810\_11663/app\_query.fasta\_1.519  
-DB=N Geneseq 16Dec04 -OFT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09380203@CN\_1\_1\_644@runat\_13092005\_170810\_11663 -NCFU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2034	100.0	1442	6	ABN89470
2	2034	100.0	1442	6	ABD46671
3	2034	100.0	1442	8	ABX08221
4	2034	100.0	1442	8	ABZ23236
5	2034	100.0	1442	8	AAL54224

6	2034	100.0	1442	10	ADB37519	Human neu
7	2034	100.0	1442	10	ACC84453	AD7C-NTP
8	2034	100.0	1442	13	ADR14408	Human NF-
9	1787	87.9	1418	2	AAT27738	Neural th
10	1517	74.6	1381	2	AAQ77883	Neural th
11	1517	74.6	1381	2	AAT27765	AD 10-7 h
12	964.5	47.4	75252	11	ACN44450	Human gen
13	901	44.3	33362	13	ABD32700	Human can
14	892.5	43.9	92638	6	ABO88096	Human ost
15	885	43.5	173805	10	ADL13775	Osteoarth
16	869.5	42.7	24345	13	ADS36503	Human aut
17	869.5	42.7	40947	13	ADS36474	Human aut
18	863.5	42.5	110096	6	ABN95044	Gene #154
19	861.5	42.4	21666	9	ADA02702	Human Nup
20	861.5	42.4	21666	10	ADB72440	Human CA
21	861.5	42.4	21666	10	ADE95950	Human hCG
22	858	42.2	126001	12	ADH77123	Human PAZ
23	857.5	42.2	24167	5	ABA16132	Human ner
24	857.5	42.2	59999	13	ADR28249	Human low
25	857	42.1	32865	11	ACN44490	Human gen
26	852	41.9	167932	10	ADL13501	Osteoarth
27	851.5	41.9	3379	10	ADH69278	Human ocp
28	851.5	41.9	3379	12	ADF91770	Human G-p
29	848.5	41.7	15783	4	AAK90159	Human dig
30	848.5	41.7	15783	5	AAS39803	Genomic s
31	848.5	41.7	15783	9	ADB32763	Human nov
32	848	41.7	61103	11	ACN43886	Human gen
33	846	41.6	124987	12	ADP03055	Human hou
34	846	41.6	124990	13	ADS88553	Human hou
35	842.5	41.4	173805	10	ADL13775	Osteoarth
36	841.5	41.4	40645	8	ABX61804	Genomic D
37	841.5	41.4	40645	9	ADA49717	Genomic D
38	841	41.3	190117	10	ADL13780	Osteoarth
39	840.5	41.3	50000	10	ADC58843	Human IKB
40	840	41.3	27148	4	AAS28612	Genomic s
41	840	41.3	27148	10	ADG41808	Human res
42	840	41.3	27148	11	ADI97582	Human res
43	840	41.3	74234	11	ACN44594	Human gen
44	840	41.3	79544	12	ADQ97764	Human can
45	839	41.2	73063	12	ADQ97727	Human can

ALIGNMENTS

RESULT 1  
ABN89470  
ID ABN89470 standard; DNA; 1442 BP.  
XX  
AC ABN89470;  
XX  
02-SEP-2002 (first entry)  
XX  
DE Neural thread protein (NTP) encoding nucleotide sequence.  
XX  
KW Neural thread protein; NTP; Harilil peptide; Alzheimer's disease;  
KW Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;  
KW hypoxia; ischaemia; cerebral infarction; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 15..1142  
FT /tag=a  
FT /product="Neural thread protein"  
XX  
PN WO200234915-A2.  
XX  
PD 02-MAY-2002.  
XX  
PF 25-OCT-2001; 2001WO-US042813.  
XX  
PR 27-OCT-2000; 2000US-00697590.  
XX

PA (NYMO-) NYMOX PHARM CORP.  
 XX Fitzpatrick J, Averback P, Focht MSS, Bibiano R;  
 PI WPI; 2002-507998/54.  
 XX P-PSDB; AB881538.  
 DR  
 XX New Harilil peptide sequences of the Neural Thread Protein, useful in  
 PT therapeutic assays, e.g. as targets for developing drugs for treating  
 PT Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing  
 PT these diseases.  
 PT  
 XX Example 1; Fig 1; 53pp; English.  
 PS  
 XX The present invention describes a neural thread protein (NTP) peptide  
 CC having an amino acid sequence selected from AB881511 to AB881529 and  
 CC their homologues, which are referred collectively as Harilil peptides (I).  
 CC (I) have neuroprotective, nootropic, vasotropic and cerebroprotective  
 CC activities, and can be used in peptide therapy. The Harilil peptide  
 CC sequences can be used as analogues for NTP in therapeutic or diagnostic  
 CC assays by replacing NTP with the peptide in such an assay. The Harilil  
 CC peptides are also useful as a trap material in a diagnostic or  
 CC therapeutic assay. Therefore, the Harilil peptides are useful in binding  
 CC assays, protein and antibody purification, therapeutics or diagnostics.  
 CC In particular, the peptides are also useful for diagnosing Alzheimer's  
 CC disease, Down's syndrome, neuroectodermal tumours, astrocytomas,  
 CC glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides  
 CC are also useful as targets for drug development for the treatment of  
 CC these diseases. The present sequence encodes a neural thread protein  
 CC given in the exemplification of the present invention  
 XX  
 SQ Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 7.7e-149 Length: 1442  
 Score: 2034.00 Matches: 375  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-380-203-2 (1-375) x ABN89470 (1-1442)  
 QY 1 MetGluPheSerLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20  
 DB 15 ATGGAGTTTTCGCTCTGTGGCCAGCTGGAGTGCATGGCGCAATCTCAGCTCACC 74  
 QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40  
 DB 75 AACCTCCGCTCCCGGGTTCAGCGATTCTCTGCTCAGCTCCCGCTCAGCTGGATT 134  
 QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60  
 DB 135 ACAGGATGTGCACCCACCGCTCGGCTAAATTTTGTATTTTGTAGATGGAGTTT 194  
 QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspSerProSerValSerAla 80  
 DB 195 CTCCTATGTTGGTCAGCTGTCTCGAACTCCCGACCTCAGATGATCCCTCCGCTCGGCC 254  
 QY 81 SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys 100  
 DB 255 TCCCAAGTGTAGATACAGGACTGGCCACCATGCCGCTCTGCTGGCTAAATTTTGT 314  
 QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLeuGlnSerThr 120  
 DB 315 GGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTCAGCTCAAGCAGTCCACC 374  
 QY 121 CysLeuSerLeuProLysCysTrpAspTyrArgAlaAlaValProGlyLeuPheIle 140  
 DB 375 TGCTCAGCTCCCAAGTGTGGGATTACAGCGGTGCAGCGCTGCTGGCTTTTATT 434  
 QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160

Db 435 TTAATTTTTTAAAGACACAGGTGTCCCACTCTTACCAGGATGAAGTGCAGTGGTGTGAT 494  
 QY 161 HisSerSerLeuGlnProSerThrProGluIleHisHisProAlaSerAlaSerGln 180  
 Db 495 CACAGCTCACTGCAGCTTCAACTCTGAGATCAGCATCTCTCTGCTCAGCTCCCA 554  
 QY 181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIlePheAen 200  
 Db 555 GTAGCTGGGACCAAGACATGCACACATACCTGGCTAAATTTTATTTTATTTTAA 614  
 QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220  
 Db 615 TTTTGTGACAGAGTCTCAACTCTCTCACCAGGCTGGAGTGCAGTGGCGCAATCTTGGC 674  
 QY 221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240  
 Db 675 TCACGTGAACCTCTGCTCCCGGGTCAAGTTATTTCTCTGCTCCCGAGCTCTCTGATAGC 734  
 QY 241 TrpAspTyrArgTgProArgLeuAlaAsnPhePheValPheLeuValGluMetGly 260  
 Db 735 TGGGACTACAGCGCCACACGCTAGCTAATTTTGTATTTTGTAGTAGATGGG 794  
 QY 261 PheThrMetPheAlaArgLeuIleLeuSerGlyProCysAspLeuProAlaSerAla 280  
 Db 795 TTCACCATGTTCCGCCAGGTGTGATCTCTGGACCTTGTGATCTGCTCGCTCGGCC 854  
 QY 281 SerGlnSerAlaGlyIleThrGlyValSerHisHisAlaArgLeuIlePheAsnPheCys 300  
 Db 855 TCCCAAGTGTGGGATTACAGCGCTGAGCCACCGCGCTTATTTTAAATTTTGT 914  
 QY 301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320  
 Db 915 TTGTTTGAATGGAATCTCCTCTTACCCAGGCTGGAGTGCATGGCAATCTCGGC 974  
 QY 321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340  
 Db 975 TCACGTGAACCTCTGCTCCCGGGTCAAGCGATTCTCTGCTCAGCTCCCAAGCAGC 1034  
 QY 341 TrpAspTyrGlyHisLeuProHisProAlaAsnPheCysIlePheIleArgGly 360  
 Db 1035 TGGGATTACGGGCACCTGCCACACCCCGCTAAATTTTGTATTTTATTTAGAGCGCG 1094  
 QY 361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375  
 Db 1095 GTTTCACCAATTTGTAGCTGGTCTCAAACTCTCTGACCTCAGG 1139  
 RESULT 2  
 AAD46671  
 ID AAD46671 standard; DNA; 1442 BP.  
 XX  
 AC AAD46671;  
 DT 27-JAN-2003 (first entry)  
 XX  
 DE AD7c-NTP gene.  
 XX Neural thread protein; NTP; hyperplasia; hypertrophy; arteriosclerosis;  
 KW haemorrhoid; gene therapy; tumour; vascular disease; atherosclerosis;  
 KW inflammatory disease; nutritional deficiency disease; genetic disease;  
 KW autoimmune disease; metabolic disease; traumatic disease; intoxication;  
 KW infectious disease; congenital malformation; enzyme deficiency disease;  
 KW amyloid disease; fibrosis disease; storage disease; radiation disease;  
 KW poisoning; environmental disease; endocrine disease; protein therapy;  
 KW degenerative disease; mechanical disease; AD7c-NTP protein; gene; ds.  
 XX  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 15..1142  
 FT /tag= a  
 FT /product= "AD7c-NTP protein"  
 XX  
 PN WO200274323-A2.



XX 26-SEP-2002.

XX

XX 08-MAR-2002; 2002WO-IB001959.

XX PF

XX 08-MAR-2001; 2001US-0273957P.

XX PR

XX (AVER/) AVERBACK P.

XX PA

XX AVerback P;

XX PI

XX WPI: 2002-759864/82.

XX DR

XX P-PSDB; AAE29142.

XX

PT Treating a condition in a patient requiring removal or destruction of

PT cells, such as a benign or malignant tumor of a tissue or an inflammatory

PT disease, comprises administering a neural thread protein (NTP) or a NTP

PT gene to a mammal.

XX

XX Example 2; Fig 1; 70pp; English.

XX

CC The invention relates to a method for treating a condition in a patient

CC requiring removal or destruction of cells. The method involves

CC administering to a mammal a neural thread protein (NTP), or administering

CC to a tumour or other target cell a NTP gene, where the expression of the

CC NTP gene is induced resulting in expression of the NTP protein. The

CC method and NTP are useful for treating a condition in a patient requiring

CC removal or destruction of cells, such as a benign or malignant tumour of

CC a tissue, a hyperplasia, hypertrophy, or overgrowth of a tissue,

CC preferably tonsillar hypertrophy or prostatic hyperplasia, a virally,

CC bacterially, or parasitically altered tissue, or a malformation of a

CC tissue. Other conditions include a cosmetic modification to a tissue,

CC such as removal of unwanted facial hair, warts or unwanted fatty tissue,

CC a vascular disease, particularly atherosclerosis or arteriosclerosis,

CC haemorrhoids, or varicose veins, an inflammatory disease, autoimmune

CC disease, metabolic disease, hereditary/genetic disease, traumatic disease

CC or physical injury, nutritional deficiency disease, infectious disease,

CC congenital malformation, amyloid disease, fibrosis disease, storage

CC disease, enzyme deficiency disease, poisoning, intoxication, degenerative

CC disease, radiation disease, environmental disease, endocrine disease or

CC mechanical disease. The invention is useful in protein therapy and gene

CC therapy. The present sequence is AD7c-NTP gene

XX

SQ Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7.7e-149 Length: 1442

Score: 2034.00 Matches: 375

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-380-203-2 (1-375) x AAD46671 (1-1442)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20

Db 15 ATGGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCATGGCGCAATCTCAGCTCACC GC 74

QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValalaGlyIle 40

Db 75 AACCTCCGCCTCCCGGGTTCAAGCGATTCTCTCGCTCAGCGCTCCCCAGTAGTGGGATT 134

QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60

Db 135 ACAGGCATGTGCACCCACGCTCGCGTAAATTTTGTATTTTTTTTAGTAGAGATGGAGTTT 194

QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspProSerValSerAla 80

Db 195 CTCATGTTGTGAGCTGGTGTCTGAATCCCGACCTCAGATGATCCCTCCGTCTCGGCC 254

QY 81 SerGlnSerAlaArgTyrArgThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCys 100

KW cosmetic modification; vascular disease; atherosclerosis;  
KW arterioclerosis; haemorrhoid; varicose vein; inflammatory disease;  
KW autoimmune disease; metabolic disease; traumatic disease;  
KW physical injury; nutritional deficiency disease; infectious disease;  
KW amyloid disease; fibrosis disease; storage disease;  
KW congenital malformation; enzyme deficiency disease; poisoning;  
KW intoxication; environmental disease; radiation disease;  
KW endocrine disease; degenerative disease; mechanical disease; gene; ds.  
XX Homo sapiens.  
XX WO200297030-A2.  
XX PD 05-DEC-2002.  
XX PF 24-MAY-2002; 2002WO-CAN000759.  
XX PR 25-MAY-2001; 2001US-0293156P.  
XX PA (NYMO-) NYMOX CORP.  
XX PI Averbach PA;  
XX DR WPI; 2003-041406/03.  
XX PT Novel peptides similar in amino acid sequence to neural thread proteins  
XX PT (NTP), useful for treating unwanted cellular proliferations such as  
XX PT malignant tumors and prostatic hyperplasia.  
XX PS Disclosure; Fig 1; 78pp; English.  
XX CC The invention describes an NTP-peptide (I) comprising at least one amino  
XX CC acid sequence corresponding to part of the amino acid sequence of a  
XX CC neural thread protein, AD7C-NTP. The invention provides a method of  
XX CC treating a condition requiring removal or destruction of cells of a  
XX CC mammal comprising administering to a mammal, a therapeutic amount of (I).  
XX CC The treatment is administered to the mammal before, during or after  
XX CC surgical excision, transplantation, grafting, chemotherapy,  
XX CC immunotherapy, vaccination, thermal or electrical ablation, cryotherapy,  
XX CC laser therapy, phototherapy, gene therapy and/or radiation. The method is  
XX CC useful for treatment of benign or malignant tumour; hyperplasia,  
XX CC hypertrophy or overgrowth of tissue; virally, bacterially or  
XX CC parasitically altered tissue; malformation of tissue selected from lung,  
XX CC breast, stomach, pancreas, prostate, bladder, bone, ovary, skin, kidney,  
XX CC sinus, colon, intestine, rectum, esophagus, heart, spleen, salivary  
XX CC gland, blood, brain and its coverings, spinal cord, muscle, connective  
XX CC tissue, adrenal, parathyroid, thyroid, uterus, testis, pituitary,  
XX CC reproductive organs, liver, hair, gall bladder, eye, ear, nose, throat,  
XX CC tonsils, mouth and lymph nodes and lymphoid system; tonsillary  
XX CC hypertrophy; prostatic hyperplasia; cosmetic modification to a tissue;  
XX CC vascular disease (atherosclerosis or arteriosclerosis); haemorrhoids;  
XX CC varicose veins; inflammatory disease; autoimmune disease; metabolic  
XX CC disease; hereditary/genetic disease; traumatic disease; physical injury;  
XX CC nutritional deficiency disease; infectious disease; amyloid disease;  
XX CC fibrosis disease; storage disease; congenital malformation; enzyme  
XX CC deficiency disease; poisoning; intoxication; environmental disease;  
XX CC radiation disease; endocrine disease; degenerative disease and mechanical  
XX CC disease. This sequence encodes the human neural thread protein AD7C-NTP  
XX SQ Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 7,7e-149 Length: 1442  
Score: 2034.00 Matches: 375  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 8 Gaps: 0  
  
US-09-380-203-2 (1-375) x ABX08221 (1-1442)  
  
QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyValAlaIleSerAlaHisArg 20  
|||||

Db 15 ATGGAGTTTTCGCTCTTGTGGCCAGGCTGGAGTCAATGGCGCAATCTCAGCTCACCGC 74  
QY 21 AenLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValalaGlyIle 40  
Db 75 AACCTCCGCTCCCGGGTTCAAGCGATTCTCTCGCTCAGCGCTCCAGTAGTGGGATT 134  
QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuPhePheLeuValGluMetGluPhe 60  
Db 135 ACAGGCATGTGACCCACGCTCGGCTAATTTTGTATTTTATTTAGTAGAGATGAGATT 194  
QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspSerValSerAla 80  
Db 195 CTCCATGTTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCCCTCCGCTCGGCC 254  
QY 81 SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys 100  
Db 255 TCCCAAAGTGTAGATACAGGACTGGCCACCATGCTCCGCTCTGCTGGCTAATTTTGT 314  
QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120  
Db 315 GGTAAGAAACAGGGTTTCACTGATGTCCTCAAGCTGGTCTCTCGAGCTCAAGCAGTCCACC 374  
QY 121 CysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheIle 140  
Db 375 TGCCTCAGCTCCCAAGTGTGGGATTACAGCGTGCAGCGTGCCTGGCCTTTTATT 434  
QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160  
Db 435 TTATTTTAAAGACACAGGTGTCCACTCTTACCCAGGATGAAGTGCAGTGGTGTGAT 494  
QY 161 HisSerSerLeuGlnProSerThrProGluIleLysHisProAlaSerAlaSerGln 180  
Db 495 CACAGCTCACTGCAGCCTTCAACTCTGAGATCAAGCATCTCTCTGCTCAGCCTCCCAA 554  
QY 181 ValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIlePheIlePheAsn 200  
Db 555 GTAGCTGGGACCAAGACATGCACCCTACCTGCTGCTAATTTTATTTTATTTTAAAT 614  
QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220  
Db 615 TTTTGTAGACAGAGTCTCAACTCTCTACCCAGGCTGGAGTGCAGTGGCGCAATCTTGC 674  
QY 221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240  
Db 675 TCATCTCAACCTCTGCTCCCGGGTTCAAGTTATTTCTCTGCCCCAGAGCTCTCTGAGTAGC 734  
QY 241 TrpAspTyrArgArgProArgLeuAlaAsnPheValPheLeuValGluMetGly 260  
Db 735 TGGGACTACAGCGCGCCACCGCTAGCTAATTTTGTATTTTGTATTTAGTAGAGATGGGG 794  
QY 261 PheThrMetPheAlaArgLeuIleLeuLeuSerGlyProCysAspLeuProAlaSerAla 280  
Db 795 TTCACCATGTTCCGCGAGTGTGATCTCTGAGACCTTGTGATCTGCTCCCTCGCTCGGCC 854  
QY 281 SerGlnSerAlaGlyIleThrGlyValSerHisHisAlaArgLeuIlePheAsnPheCys 300  
Db 855 TCCCAAAGTGTGGGATTACAGCGTGCAGCCACACCGCCGCTTATTTTATTTTGT 914  
QY 301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320  
Db 915 TTGTTTGAATGGAATCTCACTCTGTATCCAGGCTGGAGTGCATGGCAATGGCAATCTCGGC 974  
QY 321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340  
Db 975 TCATCTGCAACCTCTCCCGGGTCAAGCGAATTTCTCTGTCTCAGCTCCCAAGCAGC 1034  
QY 341 TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGlyGly 360  
Db 1035 TGGGATTACGGGCACCTGCCACACACCCCGCTAATTTTGTATTTTCAATTAGAGCGGG 1094  
QY 361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375  
Db 1095 GTTTCACCATATTTGTGAGGCTGGTCTCAAACTCTGACCTCAGCTCAGG 1139



RESULT 5	
AAU54224	
ID	AAU54224 standard; DNA; 1442 BP.
XX	
AC	AAU54224;
XX	
DT	27-MAR-2003 (first entry)
XX	
DE	AD7C-neural thread protein encoding DNA.
XX	
DE	Nootropic; neuroprotective; cell death; tissue necrosis; NTP;
KW	neural thread protein; neurodegenerative disorder; Alzheimer's disease;
KW	gene; ds.
XX	
OS	Unidentified.
XX	
OS	WO200292115-A2.
XX	
PN	21-NOV-2002.
XX	
PD	16-MAY-2002; 2002WO-CA000712.
XX	
PF	16-MAY-2001; 2001US-0290971P.
XX	
PR	(NYMO-) NYMOX CORP.
XX	
PA	Averback PA;
XX	
PI	WPI; 2003-129234/12.
XX	
DR	P-PSDB; ABJ19445.
XX	
PT	Preventing and/or inhibiting cell death and/or tissue necrosis in a
PT	tissue for treating a neurodegenerative disorder, e.g. Alzheimer's
PT	disease, by contacting the live tissue with at least one segment of
PT	neural thread proteins (NTP).
XX	
PS	Disclosure; Fig 1; 60pp; English.
XX	
CC	The invention relates to a novel method for preventing and/or inhibiting
CC	cell death and/or tissue necrosis in a tissue comprising contacting the
CC	live tissue with at least one segment of neural thread proteins (NTP).
CC	The methods are composition are useful for treating a neurodegenerative
CC	disorder, such as Alzheimer's disease. This polynucleotide sequence
CC	represents the DNA encoding the AD7C-NTP protein relating to the
CC	invention
XX	
SQ	Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	7,7e-149 Length: 1442
Score:	2034.00 Matches: 375
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	Gaps: 0
US-09-380-203-2 (1-375) x AAU54224 (1-1442)	
Qy	1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
Db	15 ATGGAGTTTTCGCTCTGTGGTCCAGGCTGGGTGCAATGGCGCAATCTCAGCTCACCGC 74
Qy	21 AsnLeuArgLeuProGlySerSerAspSerProIlaSerAlaSerProValAlaGlyIle 40
Db	75 AACCTCCGGCTCCCGGGTTCAAGCGAATTTCTCTGCTCAGCTCCCGCAGTAGCTGGGATT 134
Qy	41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60
Db	135 ACAGGATGTGCACCCACCTCCGGCTAATTTGTATTTTTTTTAGTAGAGATGGAGTTT 194
Qy	61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspProSerValSerAla 80

Antiatherosclerotic; Antiarteriosclerotic; Vasotropic; Antiinflammatory; Immunosuppressive; Tranquillizer; Antileptic; Virucide; AD7C-NTP; neural thread protein; neuritic sprouting; gene; ds.

**Homo sapiens.**

```
Key      Location/Qualifiers
CDS      15..1142
          /*tag= a
          /product= "AD7c-NTP"
```

WO2003008444-A2.

30-JAN-2003.

19-JUL-2002: 2002WO-CA001106.

19-JUL-2001: 2001US-0306150P.

15-JUL-2001; 2001US-0306161P.  
16-NOV-2001: 2001US-0331477P.

(NYMO-) NYMOX CORP.

Averback PA. Gemme

WPI: 2003-248000/24

P-PSDB; ADB3/64Z.

unwanted cellular

hair, warts and unwanted fatty tissue.  
Disclosure; Fig 1; 109pp; English.

The present invention relates to A

useful for treating a condition in a patient requiring removal or destruction of cells. The condition can be selected from benign or malignant tumour of a tissue, hyperplasia, hypertrophy or overgrowth of a tissue, virally, bacterially or parasitically altered tissue, or malformation of a tissue, where the tissue is selected from lung, breast, stomach, pancreas, prostate, bladder, bone, ovary, skin, kidney, sinus, colon, intestine, stomach, rectum, oesophagus, heart, spleen, salivary gland, blood, brain and its coverings, spinal cord and its coverings, muscle, connective tissue, adrenal, parathyroid, thyroid, uterus, testis, pituitary, reproductive organs, liver, gall bladder, eye, ear, nose, throat, tonsils, mouth, lymph nodes and lymphoid tissue. The condition is preferably tonsillary hypertrophy, prostatic hyperplasia, psoriasis, eczema, dermatosis, cosmetic modification to a tissue (skin, eye, ear, nose, throat, mouth, muscle, connective, hair or breast tissue), vascular disease (atherosclerosis or arteriosclerosis), haemorrhoids, varicose veins, inflammatory disease, autoimmune disease, metabolic disease, hereditary/genetic disease, traumatic disease or physical injury, nutritional deficiency disease, infectious disease, amyloid disease, fibrosis disease, storage disease, congenital malformation, enzyme deficiency disease, poisoning, intoxication, environmental disease, radiation disease, endocrine disease, degenerative disease and mechanical disease. The peptides are useful for treating unwanted cellular proliferations, glandular (e.g. prostate) hyperplasia, unwanted facial hair, warts and unwanted fatty tissue, or for preparing antibodies that recognize and/or bind to Related proteins, Related peptides or NTP peptides. The present sequence is the coding sequence for AD7c-NTP, used to illustrate the invention. AD7c-NTP is a ~41kD membrane associated phosphoprotein with functions associated with neuritic sprouting.

Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 U; 0 Other;

**Alignment Scores:**

Pred. No.:	7.7e-149	Length:	1442
Score:	2034.00	Matches:	375
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0

DB: 10 Gaps: 0

US-09-380-203-2 (1-375) x ADB37519 (1-1442)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyValaIleSerAlaHisArg 20  
 DB 15 ATGGAGTTTTCGCTCTTGTTGCCAGGCTGGAGTGCATGGGCAATCTCAGCTCACCGC 74  
 QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValaAlaGlyIle 40  
 DB 75 AACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCTCAGCCCTCCCAAGTAGCTGGGATT 134  
 QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60  
 DB 135 ACAGGCATGTCACCCAGCTCGGCTAAATTTGTATTTTTTTTTTAGTAGAGATGAGATT 194  
 QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAla 80  
 DB 195 CTCATGTGTGTCAGGCTGGTCTCGAACTCCCGACTCAGATGATCCCTCCGCTTCGGCC 254  
 QY 81 SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys 100  
 DB 255 TCCCAAGTGTAGATACAGGACTGGCCACCATCGCGCTCTGCCTGGCTAAATTTTGT 314  
 QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuIleGlnSerThr 120  
 DB 315 GGTAGAAACAGGGTTTCACTGATGTGCCAAGCTGGTCTCTCTGAGCTCAAGCAGTCCACC 374  
 QY 121 CysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheIle 140  
 DB 375 TGCCTCAGCCTCCCAAGTGTGGATTACAGCGGTGAGCCGCTGGCCCTTTTATT 434  
 QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160  
 DB 435 TTATTTTTTAAGACACAGGTGTCCTACTCTTACCCAGGATGAAGTGCAGTGGTGTGAT 494  
 QY 161 HisSerSerLeuGlnProSerThrProGluLeuLysHisProProAlaSerAlaSerGln 180  
 DB 495 CACAGCTCAGTCGAGCCTTCAACTCTGAGATCAAGCATCTCTCTGCCTCAGCCTCCCAA 554  
 QY 181 ValAlaGlyThrLysAspMetHisIstYrThrTrpLeuIlePheIlePheIlePheAsn 200  
 DB 555 GTAGCTGGGACCAAGACATGCAACACTACACCTGGCTAAATTTTATTATTTTAAAT 614  
 QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220  
 DB 615 TTTTGTGACACAGAGTCTCAACTCTGTACCCAGGCTGGAGTGCAGTGGCGCAATCTTGGC 674  
 QY 221 SerLeuGlnProLeuProProGlyPheLysLeuPheSerCysProSerLeuLeuSerSer 240  
 DB 675 TCACGTCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCTGCCCCAGCCTCTCTGAGTAGC 734  
 QY 241 TrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeuValGluMetGly 260  
 DB 735 TGGGACTACAGCGGCCCAACAGCCTAGCTAAATTTTGTATTTTATTAGTAGAGATGGGG 794  
 QY 261 PheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeuProAlaSerAla 280  
 DB 795 TTCACCATGTTCCGACGGTTGATCTTGATCTCTGGACCTTGTGATCTGCTGCCTCGGCC 854  
 QY 281 SerGlnSerAlaGlyIleThrGlyValSerHisAlaArgLeuIlePheAsnPheCys 300  
 DB 855 TCCCAAGTGTGGGATTACAGGCGTGAGCCACCAAGCCCGCTTATTTTAAATTTTGT 914  
 QY 301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320  
 DB 915 TTGTTTGAAATGGAAATCTCACTCTGTTACCCAGGCTGGAGTGCATGGCCAAATCTCGGC 974  
 QY 321 SerLeuGlnProLeuProProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340  
 DB 975 TCATGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCTGTCTCAGCCTCCCAAGCAGC 103  
 QY 341 TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGlyVal 360

Db 1035 TGGGATTACGGGACCTGCACACACCCCGCTAAATTTTGTATTTTCAATAGAGCGGG 1094  
QY 361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375  
Db 1095 GTTTCACCATATTTGTTCAGGCTGGTCTCAAACTCCTGACCTCAGG 1139

RESULT 7  
ACC84453  
ID ACC84453 standard; DNA; 1442 BP.  
XX ACC84453;  
AC ACC84453;  
XX 28-AUG-2003 (first entry)  
XX AD7c-NTP encoding sequence.  
XX Cytostatic; Antibacterial; Immunosuppressive; Antiinflammatory;  
KW neural thread protein; NTP; tumour; ds.  
KW Unidentified.  
OS  
XX WO2003008443-A2.  
XX 30-JAN-2003.  
XX 19-JUL-2002; 2002WO-CA001105.  
XX 19-JUL-2001; 2001US-0306150P.  
PR 19-JUL-2001; 2001US-0306150P.  
PR 16-NOV-2001; 2001US-0331477P.  
XX (NYMO-) NYMOX CORP.  
PA  
XX Averbach PA;  
XX WPI; 2003-2479999/24.  
XX Novel neural thread protein peptide, referred as cell death peptide,  
PT useful for treating prostatic hyperplasia, psoriasis, eczema, dermatosis,  
PT atherosclerosis, cosmetic modification to skin, throat, mouth, muscle.  
XX Disclosure; Fig 1; 77pp; English.

CC The present invention relates to a neural thread protein (NTP) peptide  
CC referred to as cell death peptide. Thought to be cytostatic,  
CC antibacterial, immunosuppressive and antiinflammatory. It is useful for  
CC treating a condition in a patient requiring removal or destruction of  
CC cells, for treating a condition such as benign or malignant tumor,  
CC inflammatory disease, autoimmune disease and infectious disease. The  
CC peptide useful for treatment is derived from the amino acid sequence for  
CC a pancreatic thread protein. The peptide is conjugated, linked or bound  
CC to a molecule chosen from antibody or its fragment, antibody-like binding  
CC molecule, where the molecule has a higher affinity for binding to a tumor  
CC or other target than binding to other cells. Treatment using NTP peptides  
CC can remove benign tumors with less risk and fewer of the undesirable side  
CC effects of surgery. The present sequence is an NTP encoding sequence  
XX  
SQ Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 7.7e-149 Length: 1442  
Score: 2034.00 Matches: 375  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0  
US-09-380-203-2 (1-375) x ACC84453 (1-1442)  
QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlnValAlaIleSerAlaHisArg 20  
Db 15 ATGGAGTTTTCGCTCTTGTGGCCAGGCTGGAGTGCATGGCGCAATCTCAGCTCACCGC 74

QY 21 AnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40  
Db 75 AACCTCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGGCTCCCGAGTAGCTGGATT 134  
QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60  
Db 135 ACAGGCATGTGCACCCACGCTCGGCTAAATTTTGTATTTTATTTAGTAGAGATGGAGTTT 194  
QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAla 80  
Db 195 CTCCATGTTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCCCTCGCTCGGCC 254  
QY 81 SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys 100  
Db 255 TCCCAAAAGTGTAGATACAGGACTGGCCACCATGCGGCTCTGCTGCTGCTAAATTTTGT 314  
QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120  
Db 315 GGTAGAAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTCTGAGCTCAAGCAGTCCACC 374  
QY 121 CysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheIle 140  
Db 375 TGCCTCAGCCTCCCAAGTGTGGATTTACAGGCGTGCAGCGTGCCTGGCCTTTTATT 434  
QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160  
Db 435 TTATTTTTTTTAAAGACACAGGTGTCCCACTCTTACCCAGGATGAAGTGCAGTGTGTGAT 494  
QY 161 HisSerSerLeuGlnProSerThrProGluIleLysHisProProAlaSerAlaSerGln 180  
Db 495 CACAGCTCACTGCAGCCTTCAACTCTGAGATCAAGCATCTCTCTGCTGCTGCTGCTGCCAA 554  
QY 181 ValAlaGlyThrLysAspMetHisTyrThrTrpLeuIlePheIlePheIlePheAsn 200  
Db 555 GTAGCTGGGACCAAGACATGCACCCTACCTGCTAAATTTTGTATTTTATTTTAAAT 614  
QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220  
Db 615 TTTTGTAGACAGAGTCTCAACTCTGTCCAGGCTGGAGTGCAGTGGCGCAATCTTGGC 674  
QY 221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerSerLeuLeuSerSer 240  
Db 675 TCACGTGCAACCTCTCTGCTCCCGGTTCAAGTTATTTCTCTGCCCCAGGCTCTCTGAGTAGC 734  
QY 241 TrpAspTyrArgArgProArgLeuAlaAsnPhePheValPheLeuValGluMetGly 260  
Db 735 TGGGACTACAGCGCGCCACCCAGCTAGCTAAATTTTGTATTTTGTATTTAGTAGAGATGGG 794  
QY 261 PheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeuProAlaSerAla 280  
Db 795 TTCACATGTTGGCCAGGTTGATCTTGATCTCTGGACCTTGTGATCTGCTGCTGCTGCGCC 854  
QY 281 SerGlnSerAlaGlyIleThrGlyValSerHisAlaArgLeuIlePheAsnPheCys 300  
Db 855 TCCCAAAAGTGTGGAGTTACAGGCGTGAGCACACCGCCGCTTATTTTAAATTTTGT 914  
QY 301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320  
Db 915 TTGTTTGAATGGAATCTCACTCTCTTACCAGGCTGGAGTGCATGGCAATGGCCAAATCTCGGC 974  
QY 321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340  
Db 975 TCACGTGCAACCTCTGCTCCCGGCTCAAGCGATTCTCTGCTGCTCAGGCTTCCCAAGCAGC 1034  
QY 341 TrpAspTyrGlyHisLeuProHisProAlaAsnPheCysIlePheIleArgGlyGly 360  
Db 1035 TGGGATTTACGGGCACCTGCCACCAACCCCGCTAAATTTTGTATTTTCAATTAGAGCGGG 1094  
QY 361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375  
Db 1095 GTTTCACCATATTTGTACAGGCTGGTCTCAAACTCTCTGACCTCAGG 1139

## RESULT 8

ADRI14408  
 ID ADRI14408 standard; DNA; 1442 BP.  
 XX AC ADRI14408;  
 XX DT 21-OCT-2004 (first entry)  
 XX DE Human NF-kappaB pathway-associated gene SeqID409.  
 XX DE  
 KW NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;  
 KW antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;  
 KW antihistaminic; immunomodulator; cerebroprotective; vasotropic;  
 KW immunosuppressive; vulnary; gene therapy; immune disorder;  
 KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;  
 KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;  
 KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;  
 KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;  
 KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;  
 KW viral replication; host cell survival; evasion of immune response;  
 KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;  
 KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;  
 KW autoimmune disorder; hyper immune activity;  
 KW aberrant acute phase response; hypercongenital condition; birth defect;  
 KW necrotic lesion; wound; organ transplant rejection;  
 KW aberrant signal transduction; proliferating disorder; cancer;  
 KW HIV propagation; gene, ds; human.  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO2004065577-A2.  
 PN  
 XX  
 PD 05-AUG-2004.  
 XX  
 PF 13-JAN-2004; 2004WO-US000798.  
 XX  
 PR 14-JAN-2003; 2003US-0440068P.  
 PR 12-MAY-2003; 2003US-0469757P.  
 XX  
 PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX  
 XX Nadler SG, Neubauer MG, Feder JN, Carman J;  
 PI  
 XX  
 DR WPI: 2004-562168/54.  
 DR P-PSDB; ADRI14409.  
 XX  
 PT New isolated polynucleotides and polypeptides associated with NF-kappaB  
 PT pathway, useful for diagnosing, treating, or preventing disorders or  
 PT diseases associated with NF-kappaB pathway.  
 XX  
 PS Claim 1; SEQ ID NO 409; 237pp; English.  
 XX  
 CC This invention relates to the novel association of protein sequences (and  
 CC the genes which encode them) to the NF-kappaB pathway. The invention may  
 CC be useful for the production of compounds with an antiinflammatory,  
 CC cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,  
 CC gastrointestinal-Gen, antiasthmatic, antihistaminic, anticancer,  
 CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or  
 CC vulnary activity or for gene therapy. The proteins and nucleotides are  
 CC useful for diagnosing, preventing, treating, or ameliorating conditions  
 CC or diseases associated with the NF-kappaB pathway. The condition is an  
 CC immune disorder, an inflammatory disorder, an inflammatory disorder  
 CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,  
 CC hepatic disorders, Hodgkin's lymphoma, haematopoietic tumours, hyper-IgM  
 CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic  
 CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,  
 CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell  
 CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory  
 CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick  
 CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper  
 CC immune activity, disorders related to aberrant acute phase responses,  
 CC hypercongenital conditions, birth defects, necrotic lesions, wounds,  
 CC organ transplant rejection, conditions related to organ transplant

CC rejection, disorders related to aberrant signal transduction,  
 CC proliferating disorders, cancers and HIV propagation in cells infected  
 CC with other viruses. The present sequence is that of a human gene which is  
 CC subject to the novel association with the NF-kappaB pathway of the  
 CC invention. Note: This sequence does not appear in the specification but  
 CC was obtained by the indexer from Genbank.  
 XX

SQ Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 7.7e-149 Length: 1442  
 Score: 2034.00 Matches: 375  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-09;380-203-2 (1-375) x ADRI14408 (1-1442)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20  
 DB 15 ATGGAGTTTTCGCTCTTGTGCCCAGGCTGGAGTGCATATGGCGCAATCTCAGCTCACCOC 74  
 QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValaGlyIle 40  
 DB 75 AACCTCCGCTCCCGGGTTCAAGCGATTCTCTCGCTCAGCCTCCCGTAGCTGGATT 134  
 QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60  
 DB 135 ACAGGCATGTGCACCCAGCTCGGCTAATTTTGTATTTTATTTTATAGATGGAGTTT 194  
 QY 61 LeuHisValGlyGlnAlaGlyLeuLeuProThrSerAspSerProSerValSerAla 80  
 DB 195 CTCATGTTGTCAGGCTGGTCTCGAACTCCGACTCAGATGATCCCTCCCTCTCGGCC 254  
 QY 81 SerGlnSerAlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys 100  
 DB 255 TCCCAAGTGTAGATACAGGACTGGCCACCATGCGGCTCTGCTGGCTAATTTTGT 314  
 QY 101 GlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThr 120  
 DB 315 GGTAGAACACGGGTTTCACTGATGTGCCCAAGCTGCTCTCCAGCTCAAGCAGTCCACC 374  
 QY 121 CysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeuPheIle 140  
 DB 375 TGCTTCACCTCCCAAGTGTGGATTTACAGCGGTGACGCGTCCCTGGCTTTTATTTAT 434  
 QY 141 LeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAsp 160  
 DB 435 TTATTTTAAAGACACAGGTGTCCCACTCTTACCAGGATGAAGTGCAGTGGTGTGAT 494  
 QY 161 HisSerSerLeuGlnProSerThrProGluLeuLysHisProAlaSerAlaSerGln 180  
 DB 495 CACAGCTCACTGCAGCTTCAACTCCTGAGATCAAGCATCTCTCTGCTCCCTCAGCTCCCAA 554  
 QY 191 ValAlaGlyThrLysAspMetHisIleTyrThrTrpLeuIlePheIlePheIlePheIle 200  
 DB 555 GTAGCTGGGACCAAGACATGCACCATACACCTGGCTGGCTAAATTTTATTTTATTTTAAAT 614  
 QY 201 PheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGly 220  
 DB 615 TTTTGTAGACAGAGTCTCAACTCTGTACCAGGCTGGAGTGCATGGCGCAATCTTGGC 674  
 QY 221 SerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerSerLeuSerSer 240  
 DB 675 TCACGTCAACCTCTGCTCCCGGGTTCAAGTTAATTTCTCTGCCCGCCAGCTCTCTGAGTAGC 734  
 QY 241 TrpAspTyrArgArgProArgLeuAlaAsnPhePheValPheLeuValGluMetGly 260  
 DB 735 TGGGACTACAGCGGCCCCACCGCTAGCTAATTTTGTATTTTATTTTATAGATGGGG 794  
 QY 261 PheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeuProAlaSerAla 280



Db 795 TTCACATGTCGCCAGGTTGATCTTGATCTCTGGACCTTGATCTGCCTCGCC 854  
 QY 281 SerGlnSerAlaGlyIleThrGlyValSerHisAlaArgLeuIlePheAsnPheCys 300  
 Db 855 TCCAAAGTGTGGGATTACAGCGGTGAGCCACCGCCGGCTATTTTAAATTTTGT 914  
 QY 301 LeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGly 320  
 Db 915 TTGTTGAATGGAATCTCACCTCTGTATCCAGGCTGGAGTGAATGGCCAAATCTCGGC 974  
 QY 321 SerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSer 340  
 Db 975 TCACGTCAACCTCTGCTCCCGGCTCAAGCGATTCTCTGTCTCAGCTCCCAAGCAGC 1034  
 QY 341 TrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGlyGly 360  
 Db 1035 TGGGATTACGGGCACCTGCCACACACACCGCCGCTAATTTTGTATTTTTCATTAGAGCGGG 1094  
 QY 361 ValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375  
 Db 1095 GTTTCACCATATTTGTCAAGCTGGTCTCAAACTCTCGACCTCAGG 1139

RESULT 9

AAT27738  
 ID AAT27738 standard; DNA; 1418 BP.  
 XX AC AAT27738;  
 XX DT 13-NOV-1996 (first entry)  
 XX DE Neural thread protein coding sequence.  
 XX KW Neural thread protein; NTP; diagnosis; detection; Alzheimer's disease;  
 KW neuroectodermal tumour; malignant astrocytoma; monoclonal antibody;  
 KW binding fragment; ds.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT CDS 14..1207  
 FT /\*tag= a  
 FT /product= "Neural thread protein."  
 XX PN WO9615272-Al.  
 XX PD 23-MAY-1996.  
 XX PF 14-NOV-1995; 95WO-US017111.  
 XX PR 14-NOV-1994; 94US-00340426.  
 XX PA (GENO ) GEN HOSPITAL CORP.  
 XX PI De La Monte S, Wands JR;  
 XX DR WPI; 1996-259865/26.  
 DR P-FSDS; AAR95913.  
 XX PT Detection of neural thread protein in diagnosis of Alzheimer's disease -  
 PT also NTP DNA and protein sequences used in gene and anti:sense therapy.  
 PS Claim 24; Page 168-170; 238pp; English.

CC A method for detecting the presence of neural thread protein (NTP) having  
 CC a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human subject  
 CC comprises (a) contacting a sample from a human subject that is suspected  
 CC of containing the NTP with at least one molecule capable of binding to  
 CC the protein; and (b) detecting any of the molecule bound to the protein.  
 CC The binding molecule is selected from an antibody free of natural  
 CC impurities, a monoclonal antibody or a binding fragment of either of  
 CC these. The method may be used for diagnosing the presence of Alzheimer's  
 CC disease, neuroectodermal tumours and a malignant astrocytoma in a human

SQ Sequence 1418 BP; 302 A; 396 C; 315 G; 405 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1,29e-129 Length: 1418  
 Score: 1787.00 Matches: 363  
 Percent Similarity: 95.54% Conservative: 1  
 Best Local Similarity: 95.28% Mismatches: 11  
 Query Match: 87.86% Indels: 10  
 DB: 2 Gaps: 0

US-09-380-203-2 (1-375) x AAT27738 (1-1418)

QY 1 MetGluPheSerLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20  
 Db 14 ATGGAGTTTTCGCTCTTGTGCCCAGGCTGGAGTGCATGGCGCAATCTCAGCTCAGCCG 73  
 QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProVal-AlaGlyIle 40  
 Db 74 AACCTCCGCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCCCGAGTAGGCTGGAT 133  
 QY 40 eThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPh 60  
 Db 134 TACAGGCATGTGCAC-CACGCTCGGCTAATTTTGTATTTTATTTTGTAGATGAGATT 192  
 QY 60 eLeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspProSerValSerAl 80  
 Db 193 TCTCCATGTTGGTCAGGCTGGTCTCGAACT-CCGACCTCAGATGATCTCCCGTCTCGGC 251  
 QY 80 aSerGlnSerAlaArgTyrArgThrGlyHisAlaArg-LeuCysLeuAlaAsnPheC 100  
 Db 252 CTCCTCCAAAGTGTAGATACAGACTGAGCACCAGTCCCGGCTCTGCTGGCTAATTTT 311  
 QY 100 yGlyArgAsnArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerT 120  
 Db 312 GTGGTAGAAACAGGGTTTCACTGATGTGCCAAGTGTCTCTCAGCTCAAGCAGTCCA 371  
 QY 120 hrCysLeuSerLeuProLysCysTrpAspTyrArgAlaAlaValProGlyLeuPheI 140  
 Db 372 CCTGCTCAGCCTCCCAAAGTGTGGGATTACAGCGGTGAGCTGGAGTGCAGTGGCAATCTT 431  
 QY 140 leLeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysA 160  
 Db 432 TTTTATTTTAAAGACACAGGTGTCCCACTTTACCCAGGATCAAGTCAGTGTGTG 491  
 QY 160 spHisSerSerLeuGlnProSerThrProGluIleLysHisProProAlaSerAlaSerG 180  
 Db 492 ATCAGAGCTCACTGCAGCTTCAACTCT-GAGATCAAGCATCTCTCTGCTCAGCCTCC 550  
 QY 180 In-ValAlaGlyThrLysAspMethHisIstYrThrTrpLeuIlePheIlePhe 199  
 Db 551 AAAGTAGCTGGGACCAAGACATGCACCTACACCTGGCTAATTTTATTTTATTTT 610  
 QY 200 AsnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeu 219  
 Db 611 AATTTTGTAGACAGAGTCTCAACTCTGTCAAGGCTGGAGTGCAGTGGCGCAATCTT 670  
 QY 220 GlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSer 239  
 Db 671 GGCTCACTGCAACCTCTGCTCCCGGTTCAAGTTTATTTCTCTGCTCCCGCAGCTCTGAGT 730  
 QY 240 SerTrpAspTyrArgArgProArgLeuAlaAsnPhePheValPheLeuValGluMet 259  
 Db 731 AGCTGGGACTTACAGCGGCCCAACCGCTAGCTAATTTTGTATTTTATTTTATTTT 790  
 QY 260 -GlyPheThrMetPheAlaArgLeuIleLeu-IleSerGlyProCysAspLeuProAla 279  
 Db 791 GGGTTTCAACATGTTTCGCGAGGTTGATGTAGATCTCTTGACCTTGTGTGCTGCTGCT 850  
 QY 279 eAlaSerGlnSerAlaGlyIleThrGlyValSerHisHis-AlaArgLeuIlePheAsn 298  
 Db 851 CGGCTCCCAAGTGTGGGATTACAGGACGTGACGCCCGCCCGCCGCTTATTTTAAAT 910  
 QY 299 PheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsn 318

Db 911 TTTTGTGTTTGAATGAACTCTCACTCTGTATTCCAGGCTGGAGTGCATGGCCAAAT 970  
 QY 319 LeuGlySerLeuGlnProLeuProProGlyLeuLysArgPheSerCysLeuSerLeuPro 338  
 Db 971 CTCGGCTCACTGAACCTCTGCTCCCGGGCTCAAGCGATTCTCTGTCTCAGCTCCCA 1030  
 QY 339 SerSerTrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheLeuArg 358  
 Db 1031 ASCAGCTGGATTACGGGCACCTGCA-CCACACCCCGCTAATTTTGTATTTCATTAGA 1089  
 QY 359 GlyGlyValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeuArg 375  
 Db 1090 GCGCGGGTTTCAACCATATTGTGAGGCTGGTCTCAAACTCTCAGCTCAGG 1140

## RESULT 10

AAQ77883

ID AAQ77883 standard; cDNA; 1381 BP.

AC AAQ77883;

XX 25-MAR-2003 (revised)

DT 06-JUL-1995 (first entry)

XX Neural thread protein AD10-7 cDNA.

XX Neural thread protein AD10-7; Alzheimer's; neuroectodermal tumours;

KW malignant astrocytomas; glioblastomas; ss.

XX Rattus rattus.

OS W09423756-A1.

FN 27-OCT-1994.

PD 20-APR-1994; 94WO-US004321.

XX 20-APR-1993; 93US-00050559.

PR (GEHO) GEN HOSPITAL CORP.

PA De La Monte SM, Wands JR;

XX WPI; 1994-341497/42.

DR Detection of neural thread proteins - to detect sporadic and familial

PT Alzheimer's disease, neuroectodermal tumours, malignant astrocytomas and

PT glioblastomas (Eng).

XX Example 4; Fig 16R; 158pp; English.

PS AAQ77883 is the AD10-7 neural thread protein (NTP) cDNA. This sequence

CC was used in the development of an antibody dependent method, for the

CC detection of NTPs. This new method could be used to diagnose Alzheimer's

CC disease (differentiating between sporadic and familial), neuroectodermal

CC tumours, malignant astrocytomas and glioblastomas. (Updated on 25-MAR-

CC 2003 to correct PN field.)

XX SQ Sequence 1381 BP; 296 A; 385 C; 301 G; 399 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	1.31e-108	Length:	1381
Score:	1517.00	Matches:	343
Percent Similarity:	90.16%	Conservative:	5
Best Local Similarity:	88.88%	Mismatches:	25
Query Match:	74.58%	Indels:	22
DB:	2	Gaps:	3

US-09-380-203-2 (1-375) x AAQ77883 (1-1381)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20

Db 14 ATGGAGTTTTCGCTCTTGTGTCAGGCTGGAGTGCATGGCGCAATCTCAGCTCACCGC 73

QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40  
 Db 74 AACCTCCGCCCTCCCGGGTTCAAGCATTTCTCTGCTCAGCTCAGCTCCCGAGTAGCTGGATT 133  
 QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60  
 Db 134 ACAGGCATGTGCAC-CACGCTCGGCTAATTTTGTATTATTTTATTTAGTAGAGATGGAGTTT 192  
 QY 61 --LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerA 80  
 Db 193 AACCTCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATCTCCGCTCGCGC 252  
 QY 80 laSerGlnSerAlaArgTyrArgThrGly-----HisHisAlaArg-LeuCysLeuAla 97  
 Db 253 CTGCCCAAGTGTCTGAGATT---ACAGGCATGAGCCACCATGCCCCGGGCTCTGCTGGCT 309  
 QY 98 AsnPheCysGlyArgAsnArgValSerLeuMet-CysProSerTrpSerProGluLeu 117  
 Db 310 AATTTTGTGGTAGAAACAGGGTTTCACTGATGTTGCCCAAGCTGGTCTCTCCTCAGCTCAA 369  
 QY 117 sGlnSerThrCysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProG 137  
 Db 370 GAGTCCACCTGCTCAGCTCCCAAGTGTGGGATTACAGGCGTCA-GCCGTGCCGTGG 428  
 QY 137 yLeuPheIleLeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValG 157  
 Db 429 CTTTTTATTTATTTATTTTATTTTAAACACAGGTGTACCACTCTTACCCAGGATGAGTGA 488  
 QY 157 nTrpCysAspHisSerSerLeuGlnProSerThrProGluLeuLys-HisProAla 177  
 Db 489 GTGGTGTGATCACAGCTCACTGCAGCTTCAACTCTCAGATCAAGCAATCTCTCTGCT 548  
 QY 177 erAlaSerGlnValAlaGlyThrIysAspMetHisHisTyrThrTrpLeuIlePheIleP 197  
 Db 549 CAGCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGTA-ATTTTATTT 607  
 QY 197 heIlePheAsnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpA 217  
 Db 608 TTATTTTATTTTATTTTGTAGACAGAGTCTCAC-TCGTCAACCAGGCTGGAGTGCAGTGC 666  
 QY 217 rgAsnLeuGlySerLeuGlnProLeuProGlyPheLeuLeuPheSerCysProSerL 237  
 Db 667 GCAATCTTGGCTCACTGCACCTCTGCTCCGGGTTCAAGTTATTTCTCTGCCCCAGCC 726  
 QY 237 euLeuSerSerTrpAspTyrArgArgProArgLeuAlaAsnPhePheValPheLeu 257  
 Db 727 TCCTGAGTAGCTGGGACTACAGGCGCCACACGCTAGCTAATTTTGTATTATTTAG 786  
 QY 257 alGluMet-GlyPheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeu 276  
 Db 787 TAGAGATGGGTTTCAACATGTTGCCAGGTTGATCTTGATCTCTTGACCTTGATCTG 846  
 QY 277 ProAlaSerAla-SerGlnSerAlaGlyIleThrGly-ValSerHisHisAlaArgLeu 296  
 Db 847 CTTGCTCGGCTACCCCAAGTGTGGGATTACAGTCTGCTGACTCCAGCC---CGGCCTA 903  
 QY 296 lePheAsnPheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGln 316  
 Db 904 TTTTAAATTTTGTGTTGTTGAAATGGAATCTCTACTCTGTACCAGTCCGAGTGGAGT 963  
 QY 316 rpProAsnLeuGlySerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeu 336  
 Db 964 GGCA-AATCTCGGCTACTCGCAACTCTGCTCCCGGGTC-AAGGATTTCTCTGCTCA 1021  
 QY 336 erLeuProSerSerTrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIleP 356  
 Db 1022 GCCTCCCAAGCAGCTGGGATTACGGAC-CTGCA-CCACACCCCGCTAATTTTGTATT 1079  
 QY 356 heIleArgGlyValSerProTyrLeuSerGlyTrp-SerGlnThrProAspLeuArg 375  
 Db 1080 TCATTAGAGCGGGTTTA--CCATATTTGTGAGCTGGGTCTCAAACTCTCAGCTCAGG 1137

## RESULT 11

AAT27765  
ID AAT27765 standard; cDNA; 1381 BP.

XX AC AAT27765;

XX 14-NOV-1996 (first entry)

XX AD 10-7 human neural thread protein clone (partial sequence).

DE Neutral thread protein; NTP; diagnosis; detection; Alzheimer's disease;  
KW neuroectodermal tumour; malignant astrocytoma; monoclonal antibody;  
KW binding fragment; ds.

XX Homo sapiens.

XX WO9615272-A1.

XX 23-MAY-1996.

XX 14-NOV-1995; 95WO-US017111.

XX 14-NOV-1994; 94US-00340426.

XX (GEHO ) GEN HOSPITAL CORP.

XX De La Monte S, Wands JR;

XX WPI; 1996-259865/26.

XX Detection of neural thread protein in diagnosis of Alzheimer's disease -  
PT also NTP DNA and protein sequences used in gene and anti:sense therapy.  
XX Example 4c; Fig 16R; 238pp; English.

XX A method for detecting the presence of neural thread protein (NTP) having  
CC a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human subject  
CC comprises (a) contacting a sample from a human subject that is suspected  
CC of containing the NTP with at least one molecule capable of binding to  
CC the protein; and (b) detecting any of the molecule bound to the protein.  
CC The binding molecule is selected from an antibody free of natural  
CC impurities, a monoclonal antibody or a binding fragment of either of  
CC these. The method may be used for diagnosing the presence of Alzheimer's  
CC disease, neuroectodermal tumours and a malignant astrocytoma in a human.  
CC A number of clones of neural thread protein were isolated from healthy 17  
CC -18 week old foetal human brain (HB) 2 year old temporal lobe neocortex  
CC and end stage Alzheimer's disease (AD) cerebral cortex. See AAT27753-75

SQ Sequence 1381 BP; 295 A; 385 C; 301 G; 400 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1.31e-108 Length: 1381  
Score: 1517.00 Matches: 343  
Percent Similarity: 90.16% Conservative: 5  
Best Local Similarity: 88.86% Mismatches: 25  
Query Match: 74.58% Indels: 22  
DB: 2 Gaps: 3

US-09-380-203-2 (1-375) x AAT27765 (1-1381)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20  
DB 14 ATGGAGTTTTTCGGCTCTGTGTGCCCCAGCTGGAGTGCATGGCGCAATCTCAGCTACCCG 73  
QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerProValaGlyIle 40  
DB 74 AACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCTCAGCTCCCGAGTAGCTGGGATT 133  
QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60  
DB 134 ACAGGCATGTGCAC-CACGCTCGGCTAAATTTTGTATTTTGTATAGTAGAGTAGGAGTTT 192  
QY 61 --LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspProSerValSerA 80

Db 193 AACTCCATGTTGGTCTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATCTCCCGTCTCGCG 252  
QY 80 laSerGlnSerAlaArgTyrArgThrGly-----HisHisAlaArg-LeuCysLeuAla 97  
Db 253 CTGCCCAAAGTCTGAGATT--ACAGGCATGAGCCACCATGCGCGGCTCTGCTGGCT 309  
QY 98 AsnPheCysGlyArgAsnArgValSerLeuMet-CysProSerTrpSerProGluLeu 117  
Db 310 AATTTTGTGTAGAAACAGGGTTTCACTGATGTTGGCCCAAGCTGGTCTCTGAGACTCAA 369  
QY 117 sglnSerThrCysLeuSerLeuProLysCysTrpAspTyrArgAlaValProGln 137  
Db 370 GCAGTCCACCTGCTCAGCCTCCCAAGTGTGGATTTACAGCGCTCA-GCCGTGCCTGG 428  
QY 137 yLeuPheIleLeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGln 157  
Db 429 CCTTTTATTTTATTTTATTTTAAAGACACAGGTGTACCACCTCTTACCAGGATGAAGTGCA 488  
QY 157 nTrpCysAspHisSerSerLeuGlnProSerThrProGluIleLys-HisProProAla 177  
Db 489 GTGGTGTGATCACAGCTCACTGCAGCTTCAACTCTGAGATCAAGCAATCTCTCGCT 548  
QY 177 erAlaSerGlnValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIleP 197  
Db 549 CAGCCTCCCAAGTAGCTGGACCAAGACATGCACCATACACCTGGTA-ATTTTATTTT 607  
QY 197 heIlePheAsnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpA 217  
Db 608 TTATTTTAAATTTTGTGAGACAGAGTCTCAC-TCTGTCAACCAGGCTGGAGTGCAGTGGC 666  
QY 217 rgAsnLeuGlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerL 237  
Db 667 GCAATCTTGGCTCACTGCACACCTCTGCTCCGGGTTCAAGTATTCTCTCTCCCGCAGCC 726  
QY 237 euLeuSerSerTrpAspTyrArgArgProProArgLeuAlaAsnPhePheValPheLeu 257  
Db 727 TCCTGAGTAGCTGGAGCTACAGCGGCCCAACACCGCTAGCTAATTTTGTATTTTAG 786  
QY 257 alGluMet-GlyPheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeu 276  
Db 787 TAGAGATGGGGTTTCCACCATGTTCCAGGTTGATCTTGTATCTCTTGACCTTGTGATCTG 846  
QY 277 ProAlaSerAla-SerGlnSerAlaGlyIleThrGly-ValSerHisHisAlaArgLeu 296  
Db 847 CTGCTCTCGGCTACCCAAAGTGTGGGATTACAGGTCGTGACTCCAGC---CGGCCTA 903  
QY 296 lePheAsnPheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnT 316  
Db 904 TTTTAAATTTTGTGTTGTAATGGAAATCTCATCTGTACCAGGTCGGAGTGCAAT 963  
QY 316 tpProAsnLeuGlySerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeu 336  
Db 964 GGCA-AAATCTCGGCTACTCGCAACCTCTGCTCCCGGGTCA-AAGGATTTCTCTCTCTCA 1021  
QY 336 erLeuProSerSerTrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIleP 356  
Db 1022 GCCTCCCAAGCAGCTGGGATTACGGGAC-CTGCA-CCACACCCGCTAATTTTGTATTT 1079  
QY 356 heIleArgGlyGlyValSerProTyrLeuSerGlyTrp-SerGlnThrProAspLeuArg 375  
Db 1080 TCATTAGAGCGGGTTTA--CCATATTGTTCAGGCTGGGTCTCAAACTCTCACCTCAGG 1137  
RESULT 12  
ACN44450  
ID ACN44450 standard; DNA; 75252 BP.  
XX ACN44450;  
XX ACN44450;  
XX 18-NOV-2004 (first entry)  
XX Human genomic sequence hCG27772.  
XX

KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

OS Homo sapiens.

PN WO2003073826-A2.

XX 12-SEP-2003.

XX 28-FEB-2003; 2003WO-US006235.

XX 01-MAR-2002; 2002US-00087192.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-328604/31.

XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma comprises a nucleotide sequence.

XX Claim 1; SEQ ID NO 904; Opp; English.

XX The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published

XX SQ Sequence 75252 BP; 16571 A; 19384 C; 19541 G; 19756 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.: 2.17e-63 Length: 75252  
Score: 964.50 Matches: 252  
Percent Similarity: 52.09% Conservative: 22  
Best Local Similarity: 47.91% Mismatches: 95  
Query Match: 47.42% Indels: 161  
DB: 11 Gaps: 12

US-09-380-203-2 (1-375) x ACN44450 (1-75252)

QY	4	SerLeuLeuProArgLeuGluCysAenGlyAlaIleSerAlaHisArgLeuArg	23
DB	2398	TCGCTCTTTTGGCCAGGCTGGAGTGCATGGCGCAATCTCGCTCATGCAACCTCCGC	2457
QY	24	LeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIleThrGlyMet	43
DB	2458	CTCCGGGTTCAGCACTTCTCTGCTCAGCTCCGAGTAGTGGGATTACAGCATG	2517
QY	44	CysThrHisAlaArgLeuIleLeuThrPheLeuValGluMetGluPheLeuHisVal	63
DB	2518	CGCCACACCGCCGGCTAATTGTAT---TTTAA-GTAGAGATGGGGTTTCTCCATGTT	2573
QY	64	GlyGlnAlaGlyLeuLeuProThrSerAspProSerValSerAlaSerGlnSer	83
DB	2574	GGTCAGGCTGGTCTCGAATCGGACCTCAGGTGATCCACCTGCCTCAGCCCTCCCAAGT	2633
QY	84	AlaArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCys-----Gly	101
DB	2634	GCTAGAGCGGGTGGAGCCAC-----CGGCG-CTGGCCGGTAATTTTGTATTTTAGT	2686
QY	102	ArgAsnArgValSerLeuMetCysProSerTrp-----SerProGluLeuIysGlnSer	119
DB	2687	AGAGACGGGGTTTCGCCATGTTGGCCAGGCTGGTGTGAACCTCCTCAGGTCATCC	2746

QY	120	ThrCysLeuSerLeuProLysCysTrpAspTyrArg---ArgAlaAlaValProGlyLeu	138
DB	2747	ACCCACCTTGGCTCCCAAGTCTGGAATTACATGTGTGAGCCACCACCATGCCGGCTGCT	2806
QY	139	PheIleLeu-----	141
DB	2807	TTTTTCTCTTAACCTCTTTCTTCCCACTTTAGTTAATGCAGTGTAAAGGCTAATTCAT	2866
QY	142	-----PhePheLeuArgHis-----	146
DB	2867	TAAGTGTGTCTTAATAAATGTTATTACCCCAACACATCTTTATTAGGACACCTGTTGTAT	2926
QY	146	-----	146
DB	2927	GCCAGCACTGGGATAGAACATGAGCAATGTCTCCACCTTTCTGAAGCCTACAGTCTA	2986
QY	146	-----	146
DB	2987	GTAGGAGACAAACACAAATAACAAGTTAATAGGAATGATTATTATTACTATTTTTGAGAC	3046
QY	147	ArgCysProThrLeuThrGlnAspGluValGlnTrpCysAspHisSerSerLeuGlnPro	166
DB	3047	AGAGTCTCGCTCTGTACCAGGCTGGAGTGCAGTGCAGCAATCTCGGCTCACTGCAAGCT	3106
QY	167	SerThrProGluIle-LysHisProProAlaSerAlaSerGlnValAlaGlyThrLysAs	186
DB	3107	CCACTTCCCGGTTCCAGCCATCTCTGCTCAGCTCCGAGTAGCTGGGACTGCAGG	3166
QY	186	pMetHisIstYrThrTrpLeuIlePheIle-----	196
DB	3167	CATGCGCCACCACGCCAGCTAATTTTGTATTATTTAGTAGAAACGGGGTTTCCACCATGT	3226
QY	196	-----	196
DB	3227	TGTCAGGATGTTCTTAACCTCTTGACCTCGTGATCTGCCACCTCAGCCTCCCAAAATG	3286
QY	196	-----	196
DB	3287	CTGGGATTACAGGCATGAGCCACTGTGCTGCCGGAATATTATTATTATTATTATTA	3346
QY	197	-PheIlePheAsnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTr	216
DB	3347	TTTTATTTTATTTTGTGAGAGGAGTCTCAC-TCGTCAACCCAGGCTGGAGTGCAGTG	3405
QY	216	pArgAsnLeuGlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSe	236
DB	3406	GCATGATCTCGGCTCACTGCAAGCTCTGCTCTCGGTTTCATGCCATTTCTCTGCCTCAG	3465
QY	236	rLeuLeuSerSerTrpAspTyrArgArg---ProProArgLeuAlaAsn-PhePhe---	253
DB	3466	CCTCTGTAGTAGTGGGACTACAGACGCCGCCACACACCTGGCTTAATTTTTTTTTTTT	3525
QY	254	--ValPheLeuValGluMet-GlyPheThrMetPheAlaArgLeuIleLeuIleSerGly	272
DB	3526	ATATATTTTATAGATGGGTTTACCAGTTAGCCAGATGGTCTCGATCTCTCTCA	3585
QY	273	ProCysAspLeuProAlaSerAlaSerGlnSerAlaGlyIleThrGlyValSerHisHis	292
DB	3586	CCTCGTGATCCACCCGCTCAGCCTCCCAAGTGTGGGATTACAGGCATGAGCCACCT	3645
QY	293	AlaArg-----	294
DB	3646	GCCCGCCGACAAATTAATTAATTAATCATCATGATGAGTCTTCTAGAATGCTAAAAAAA	3705
QY	295	-----LeuIlePheAsnPheCysLeuPheGluMetGluSerHisSerValThrGlnAla	312
DB	3706	GATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT	3764
QY	313	GlyValGlnTrpProAsnLeuGlySerLeuGlnProLeuProGlyLeuIysArgPhe	332
DB	3765	GGAGTGCAATGTTGGCATCTCAGCTCACTGCAACCTCCACCTCTCTGGGTTTGGAGTTC	3824

QY 333 SerCysLeuSerLeuProSerSerTrpAspTyrGlyHisLeuProHisProAlaAsn 352  
 Db 3825 TCTCCCTCAGCTCCCGAGTAGCTGGATTACAGGTGTGTGCCACCACTGGCTAAT 3884  
 QY 353 PheCysIlePheIleArg--GlyGlyValSerProTyrLeuSerGlyTyrSerGlnThr 371  
 Db 3885 TTTTGTATTTTAGTAGACAGCGGGAGTTTACCATATGGCCAGGCTGGTCTTGACT 3944  
 QY 372 ProAspLeuArg 375  
 Db 3945 CCTGACCTCAGG 3956

RESULT 13  
 ABD32700/c  
 ID ABD32700 standard; DNA; 33362 BP.  
 XX  
 AC ABD32700;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Human cancer-associated genomic DNA HD14-035.  
 XX  
 KW Human; ds; cancer-associated protein; gene; cytostatic; cancer;  
 KW Leukaemia; lymphoma; CAP.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO2004074320-A2.  
 XX  
 PD 02-SEP-2004.  
 XX  
 PF 17-FEB-2004; 2004WO-US004730.  
 XX  
 PR 14-FEB-2003; 2003US-00367094.  
 PR 14-MAR-2003; 2003US-00388838.  
 PR 15-APR-2003; 2003US-00417375.  
 PR 13-JUN-2003; 2003US-00461862.  
 PR 15-SEP-2003; 2003US-00663431.  
 PR 15-DEC-2003; 2003US-00737318.  
 XX  
 PA (SAGR-) SAGRES DISCOVERY INC.  
 XX  
 PI Morris DW, Morris DW, Malandro MS;  
 XX  
 DR WPI; 2004-652914/63.  
 XX  
 PT New isolated cancer-associated polynucleotides and polypeptides useful  
 PT for diagnosing, preventing or treating cancers, especially lymphoma and  
 PT leukemia, or in screening for agents that modulate cancer.  
 XX  
 PS claim 16; seqid 265; 310pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising at least 10  
 CC contiguous nucleotides of any of the 233 polynucleotide sequences given  
 CC in the specification, or its complement. The nucleic acids encode cancer-  
 CC associated proteins. Also included are an expression vector comprising  
 CC the isolated nucleic acid cited above, a host cell comprising the above  
 CC recombinant nucleic acid or expression vector, a microarray for detecting  
 CC a cancer-associated (CA) nucleic acid comprising at least one probe  
 CC comprising at least 10 contiguous nucleotides of any of the above-  
 CC mentioned nucleotide sequences, an isolated polypeptide (encoded within  
 CC an open reading frame of a CA sequence selected from any of the 95  
 CC polynucleotide sequences as mentioned in the specification, or its  
 CC complement), an isolated antibody, (or its antigen binding fragment) that  
 CC binds to the above polypeptide, a hybridoma that produces the above  
 CC monoclonal antibody, a pharmaceutical composition comprising the above  
 CC antibody and a pharmaceutical excipient, a kit for detecting cancer  
 CC cells (comprising the antibody cited above, methods for diagnosing cancer  
 CC or for detecting the presence or absence of cancer cells in an  
 CC individual, a method for inhibiting growth of cancer cells in an  
 CC individual, a method for delivering a therapeutic agent to cancer cells  
 CC in an individual, an electronic library comprising the above  
 CC polynucleotide or polypeptide (or their fragments), methods of screening

CC for anticancer activity or for a bioactive agent capable of modulating  
 CC the activity of a CA protein (CAP), methods for detecting cancer  
 CC associated with expression of a polypeptide in a test cell sample, a  
 CC method for treating cancers and a method for inhibiting the expression of  
 CC CA gene in a cell. The composition and methods are useful for detecting,  
 CC diagnosing, preventing and treating cancers, especially lymphoma and  
 CC leukaemia. These may also be used in screening for agents that modulate  
 CC cancer. The present sequence is a human CAP genomic sequence. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 33362 BP; 8042 A; 8668 C; 8695 G; 7957 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 6.8e-59 Length: 33362  
 Score: 901.00 Matches: 236  
 Percent Similarity: 52.97% Conservative: 23  
 Best Local Similarity: 48.26% Mismatches: 101  
 Query Match: 44.30% Indels: 134  
 DB: 13 Gaps: 9

US-09-380-203-2 (1-375) x ABD32700 (1-33362)

QY 6 LeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArgAsnLeuArgLeuPro 25  
 Db 12760 CTGTTGCCAGCGCTGGAGTGTGTGATCTTGGCTCACCGCAAGCTCCACCTCCCG 12701  
 QY 26 GlySerSerAspSerProAlaSerAlaSerProValAlaGlyIleThrGlyMetCysThr 45  
 Db 12700 GGTTCACGCCATTCTCCTGCCTCAGCTCCCGAGTAGTGGAGCTAGAGCGCTGCCAC 12641  
 QY 46 HisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPheLeuHisValGlyGln 65  
 Db 12640 CATGCCAGCTAAATTTTGTGATTTTATTAGACAGAGGGTTTCAACCATCTTAGCCAG 12581  
 QY 66 AlaGlyLeuGluLeuProThr----- 72  
 Db 12580 GATGTCCTTGATCTCTCAGCTCGTGATCTGCCGCGCTCTAAAGTGTGGGAT 12521  
 QY 72 ----- 72  
 Db 12520 TACAGCGGTGAGCCACAGTGTCCGGCCTATTATTATTATTTTATAGACAGATCTT 12461  
 QY 72 ----- 72  
 Db 12460 ACTTGTGTTGCCAGGTTGTAGTGCAGTGGCATAAATCTTGGCTCACTGCAACCTCTGCCTC 12401  
 QY 73 -----SerAspAspProSerValSerAlaSerGlnSerAlaArgTyrArgThrGlyHis 90  
 Db 12400 CTTGGTTCAAGTGATTTCTCTCGCTCAGCTCCCAAGTAGCTGGGATACAGGCACCTAC 12341  
 QY 91 HisAlaArgLeuCysLeuAlaAsnPheCys-----GlyArgAsnArgValSerLeuMet 108  
 Db 12340 CACCA-----TGCTAGTAATTTTGTATTTTATAGATAGATGGGTTTCGCCATGT 12288  
 QY 109 CysProSerTrpSer-----ProGluLeuLys-GlnSerThrCysLeuSerLeuProLys 126  
 Db 12287 TGGCAAGGCTGTGCTTGAACCTCTCAGCTCAGGTGATATACCCCACTCGGCTCCCAA 12228  
 QY 126 sCysTrpAspTyrArgAlaAlaValProGlyLeuPheIleLeuPheLeuArgHis 146  
 Db 12227 GTGTGGGATTTACAGGTGTGACCACTGGCCAGCCCACTCTCTTTTATTATAGACA 12168  
 QY 146 sArgCysProThrLeuThrGlnAspGluValGlnTrpCysAspHisSerSerLeuGlnPr 166  
 Db 12167 -GAGTCCTGCTCTGTCGCCAGCGTGGAGTGCAGTGGCATGATCTCGGCTCACTGCAATC 12109  
 QY 166 oSerThrProGluIleLys-HisProProAlaSerAlaSerGlnValAlaGlyThrLysA 186  
 Db 12108 TCCACCTCTGGGTTCAAGCTATTCTCTGCTCAGCTCCCGAGTAGCTGGGATTACAG 12049  
 QY 186 spMetHisHisTyrThrTrpLeuIlePheIlePheAsnPheLeuArgGlnSerL 206

```

Db 12048 GCATGCACCACCGCTGGCTGATTTTATT-----TTGACAGCGAGTC 12004
Qy 206 euAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGlySerLeuGlnProLeuP 226
Db 12003 TCAC-TCCTGACCCAGGCTGGAGTGCATGGCGTATCTTTGCTCACTACAACTCCGC 11945
Qy 226 roProGlyPheLysLeuPheSerCysProSerLeuLeuSerSerTrpAspTyrArg- 245
Db 11944 CGCCGGATTCAAGCATTTCTTCGCTCAGCTCCGAGTAGCTGGGATTACAGGCATG 11885
Qy 246 --ProProArgLeuAlaAsnPhePheValPheLeuValGluMet-GlyPheThrMetPhe 264
Db 11884 TGCCACCATGCGCGGCTAATT--TTTGATTTTATAGATGGGGTTTCCACCATGTTG 11827
Qy 265 AlaArgLeuLeuLeu----- 269
Db 11826 GAAGGCTGGCTTTGAACCTCGACCTCAGGTGATACCCCACTCGGCTCCCAAAGTG 11767
Qy 269 ----- 269
Db 11766 CTGGGATTACAGGTGACCACTGCGCCAGCCCTCTCTTTTAAAGACAG 11707
Qy 270 -----IleSerGlyProCysAspLeu 276
Db 11706 AGTCCTGCTCTGCGCCAGGCTGGAGTGCAGTGGCATGATCTCGCTCACTGCAACCTC 11647
Qy 277 -----ProAlaSerAlaSerGlnSerAlaGlyIleThrGly 288
Db 11646 CACCTCTTGGGTCAAGCTATTCTCTGCTCAGCTCTGAGTAGCTGGGATTACAGGC 11587
Qy 289 ValSerHisAlaArgLeuLeuPheAsnPheCysLeuPheGluMetGluSerHisSer 308
Db 11586 ATGCACCAACCGCTGGCTGATTTTAT-----TTTGACAGGAGTCTGCTCT 11537
Qy 309 ValThrGlnAlaGlyValGlnTrpProAsnLeuGlySerLeuGlnProLeuProGly 328
Db 11536 GTCAACCGCTGGAGTGCAGTGGCATGATCTTGGCTCACTACAACCTCCGCCACCCGGA 11477
Qy 329 LeuLysArgPheSerCysLeuSerLeuProSerSerTrpAspTyrGlyHisLeuProPro 348
Db 11476 TTCAAGCGGATTTCTTGCCTCAGCTCCGAGTAGCTGGGATTACAGGCACGTCACCA 11417
Qy 349 HisProAlaAsnPheCysIlePheIleArgGlyGlyValSerProTyrLeuSerGlyTrp 368
Db 11416 TCCCTGCTAATTTTGTATTTTATAGATGGGGTTTACCATTAGGCCAGGCTGG 11357
Qy 369 SerGlnThrProAspLeuArg 375
Db 11356 TCTCAAACTCCTGACCTCAGA 11336

RESULT 14
ID ABQ88096/c
XX ABQ88096 standard; cDNA; 92638 BP.
AC ABQ88096;
XX
XX
DT 18-SEP-2002 (first entry)
XX
DE Human osteoblast differentiation related cDNA SEQ ID NO 3.
XX
KW Human; osteoblast; stem cell differentiation; bone tissue deposition;
KW osteoporosis; osteopathic; ss.
XX
OS Homo sapiens.
XX
PN WO200250301-A2.
XX
PD 27-JUN-2002.
XX
XX 18-DEC-2001; 2001WO-US048276.
XX
XX 18-DEC-2000; 2000US-0255882P.
XX

```

```

PR 24-APR-2001; 2001US-0285691P.
XX (GENE-) GENE LOGIC INC.
PA (PROC ) PROCTER & GAMBLE CO.
XX
PI Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
XX Mertz L;
XX WPI; 2002-557663/59.
XX
PT Use of genes and their expression profiles associated with osteoblast
PT differentiation for screening modulators bone formation, for diagnosing
PT or treating e.g. osteoporosis, or as markers for the differentiation
PT process.
XX
PS Claim 1; SEQ ID NO 3; 78pp + Sequence Listing; English.
XX
CC The invention relates to genes and their expression profiles are used
CC for: (a) screening modulators of precursor stem cell differentiation into
CC osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
CC deposition of bone tissue, abnormal rate of osteoblast formation or
CC osteoporosis; or (c) treating or monitoring treatment of the conditions
CC cited in (b), or monitoring the progression of bone tissue deposition.
CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
CC induced abnormalities in bone formation or bone loss, conditions that
CC involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
CC or fibrous dysplasia. The present sequence is that of an osteoblast
CC differentiation associated cDNA marker of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 92638 BP; 21460 A; 24572 C; 25235 G; 21371 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,14e-57 Length: 92638
Score: 892.50 Matches: 239
Percent Similarity: 43.48% Conservative: 21
Best Local Similarity: 39.97% Mismatches: 99
Query Match: 43.88% Indels: 242
DB: 6 Gaps: 10

US-09-380-203-2 (1-375) x ABQ88096 (1-92638)
Qy 4 SerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArgAsnLeuArg 23
Db 58335 TCACCTCTTGTGCGCCAGGCTGCAGTGCATGATGCTGGCTCACTGCAACCTCTGC 58276
Qy 24 LeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIleThrGlyMet 43
Db 58275 TTCTCTGGGTTCAAGTGATCTCTCGCTCAGCTCCCAAGTAGTTGGAAATTACAGCGCC 58216
Qy 44 CysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPheLeuHisVal 63
Db 58215 TGCCACACAGCCCGACCTAATTTTGTATTTT---GTAGACAGAGGGTTTCCACCATGTT 58159
Qy 64 GlyGlnAlaGlyLeuGluLeuProThrSerAspSerProSerValSerAlaSerGlnSer 83
Db 58158 GGCCAGCGGGGCTTGAACCTCTGACCTCAGCTGATGCCCGCTCGGCTCCCAAAAT 58099
Qy 84 Ala-----ArgTyrArgThrGlyHisAlaArgLeuCysLeuAlaAsn 98
Db 58098 GCCCGGATTACAGACATGAGCTACCGT-----GCCAGCCTAAT 58060
Qy 99 Phe-----CysGlyArgAsnArgValSerLeuMetCysProSerTrpSer-----Pro 114
Db 58059 TTTTATTATTTTATAGACAGAGGTTTACCATGTTGGCCAGGCTGGTCTGAAACTCT 58000
Qy 115 GluLeuLysGlnSerThrCysLeuSerLeuProLysCysTrpAspTyrArg---ArgAla 133
Db 57999 GACTTCAGGTGATCCCGCCCACTCAGCTCCCAAGTCTGGGATTACAGGCATGAGCCA 57940

```

QY	134	AlaValProGlyPheLeuPhePheLeuArgHisArgCysProThrLeuThrGln	153	QY	294	-----ArgLeuIlePhe	297
Db	57939	CMATGCCCGCAAT-----TTTTGAGACAGGCTCTC-TCTGTCACTCAG	57893	Db	56874	AAGAATTGGGAATTATGTTTACCTCTTAAGGCAGAAATATCTACACAGGTTTTTGT	56815
QY	154	AspGluValGlnTrpCysAspHisSerLeuGlnProSerThrProGluIleLys-Hi	173	QY	297	eAsnPheCysLeuPheGluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpPr	317
Db	57892	GCTGGAGTGCAGTGGCCAGTCATAGCTCACTGCAGCCTCGACCTCAGGGCTCAAGCA	57833	Db	56814	TTTTTTTTTCTTTTGTAGATGGAGTCTCGCTCTGTTGCCAGGCTGGAGTGCAGTGCA	56755
QY	173	sProProAlaSerAlaSerGlnValAlaGlyThrLysAspMetHisHisThrTriple	193	QY	317	oAsnLeuGlySerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeuSerIle	337
Db	57832	TCCTCTGCTCAACCTCCAGTAGCTGGAGTACAGG-CAGCACCACTATGGCTGGCT	57774	Db	56754	CAATCTTGCTCACTGAAAGCTCTGCTCCCGGGTTCAACAATTTCTCTGCTCAGCT	56695
QY	193	uilePheIlePhePheAsn-----PheLeuArgGlnSerLeuAsnSe	208	QY	337	uProSerSerTrpAspTyrGlyHisLeuProProHisProAlaAsnPheCysIlePheTrl	357
Db	57773	AAATTTTACATTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	57715	Db	56694	CCGAGTAGCTGGAGTACAGGCGCCGCCACCATGTCCAGTAATTTTGTATTTTAG	56635
QY	208	rValThrGlnAlaGlyValGlnTrpArgAsnLeuGlySerLeuGlnProLeuProGln	228	QY	357	eArgGlyGlyValSerProTyrLeuSerGlyTrpSerGlnThrProAspLeu	374
Db	57714	TGTCACCCAGGCTGAAGTACAGTGGTGGCATCTTGGCTCACTGCCAGCTCCACCTCCCG	57655	Db	56634	TAGACCGGGTTTACCTTGTAGCCAGGATGGTCTCGGACTCTGACCTT	56583
QY	228	yPheLysLeuPheSerCysProSerLeuLeuSerSerTrpAspTyrArgProPro--	247	RESULT 15			
Db	57654	GTTACGCCATTCTCTGCTCAGCTCCGAGTAGCTGGACTACAGGCGCCCACTC	57595	ID	ADL13775	standard; DNA; 173805 BP.	
QY	248	-ArgLeuAlaAsnPhePheValPheLeuValGluMetGlyPheThrMetPheAlaArgLe	267	XX	ADL13775;		
Db	57594	AGCCCGAGCTAACTTTTGTATTTTGTAGAGACGGGGTTTCCACCATTTAGCCAGGAT	57535	XX	06-MAY-2004 (first entry)		
QY	267	uileLeuIleSerGlyProCysAspLeuProAlaSerAlaSerGlnSerAlaGlyIlePh	287	DE	Osteoarthritis-associated polymorphic nucleotide #307.		
Db	57534	GGTCTCGATCTCTGACCTCGGATCCGCGCTCGGCTCCAAAGTCTGGAGTAC	57475	KW	ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;		
QY	287	rglyValSerHisHisAla-----	293	KW	joint space narrowing; osteophyte development; joint pain;		
Db	57474	AGGAGTGAGCCACTATGCCCCGCTAAGTTTTTAATTTTTCATAGACACAGGTCTCACT	57415	XX	osteoarthritis; SNP; single nucleotide polymorphism.		
QY	293	-----	293	OS	Homo sapiens.		
Db	57414	ATTGTTGCCAGGCTGGTCTCAAACTTACGTGATCCTCTACCTTAGCCTCCCAAAGGCG	57355	PN	WO2003054166-A2.		
QY	293	-----	293	PD	03-JUL-2003.		
Db	57354	TGGATTACAGGTGTGGTCTCCCTATTATTAAACACCTTAGTATGTTGTCACAA	57295	PF	19-DEC-2002; 2002WO-US041225.		
QY	293	-----	293	PR	20-DEC-2001; 2001US-0342603P.		
Db	57294	TGAATCATTTAACTAACTCCATATCTTTTATTTTCACTAGATTTTCCCTTAATGT	57235	XX	(INCY-) INCYTE GENOMICS INC.		
QY	293	-----	293	PI	Jones KA, Schafer A;		
Db	57234	CCTTTTCCCTCAGAAATCCATTTCAATCATCCTGACTCTTGGATGTATGTTCTC	57175	XX	WPI; 2003-559141/52.		
QY	293	-----	293	DR	Determining susceptibility of an individual to joint space narrowing,		
Db	57174	AGACCTTCCCTGTTTGTAGACCTTACAGTTTTGGGAGTACCGGGTAGGATATGGTA	57115	XX	osteoophyte development and/or joint pain comprises identifying whether		
QY	293	-----	293	CC	the individual has at least one polymorphism in a polynucleotide encoding		
Db	57114	GAATCCCTCTATCAGGATTTGTCGTGATTTTCTCTCATGATTAGACTGAGGTCTGGGT	57055	CC	Disclosure; SEQ ID NO 307; 297pp; English.		
QY	293	-----	293	CC	The invention relates to a method of determining susceptibility of an		
Db	57054	TTTGGAGAGGAACACAGAGGTCAAGTGCCATTCTTGCCACATCTTATGAGGGTTTCATG	56995	CC	individual to joint space narrowing and/or osteophyte development and/or		
QY	293	-----	293	CC	joint pain comprising identifying whether the individual has at least one		
Db	56994	CTGACTTTGATCACCTGACTGAGTTAACTTTGCCAGACTTCTCCATTGTAAGTTACTC	56935	CC	polymorphism in a polynucleotide encoding at least one of the protein		
QY	293	-----	293	CC	listed in the specification. The methods, composition and agent are		
Db	56934	TTTTTCTTCTTCTGACTGGACCTTTTGAAGGAAGTCACTGTGTGACGCCACATTT	56875	CC	useful for modulating the susceptibility of an individual to joint space		
				CC	narrowing and/or osteophyte development and/or joint pain that is		
				CC	associated with a disease, preferably osteoarthritis. The cell line and		
				CC	the non-human animal are useful for screening for an agent for diagnosing		
				CC	an individual having susceptibility to joint space narrowing and/or		
				CC	osteoophyte development and/or joint pain. This sequence corresponds to		
				CC	the polynucleotide encoding a protein listed in the specification. (Note:		
				CC	The sequence data for this patent did not form part of the printed		
				CC	specification but was obtained in electronic format directly from WIPO at		
				CC	ftp.wipo.int/pub/published_pct_sequences).		
				XX			



SQ Sequence 173805 BP; 47663 A; 41986 C; 39245 G; 44910 T; 0 U; 1 Other;

**Alignment Scores:**

Pred. No.:	9,71e-57	Length:	173805
Score:	885.00	Matches:	236
Percent Similarity:	51.86%	Conservative:	29
Best Local Similarity:	46.18%	Mismatches:	101
Query Match:	43.51%	Indels:	147
DB:	10	Gaps:	12

US-09-380-203-2 (1-375) x ADL13775 (1-173805)

Qy	4	SerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIalleSerAlaHisArgAsnLeuArg	23
Db	77347	TCACCCCTTGTTGCCAGGCTGGAGTGCATATGGCGTGTCTCGGCTCATGGCAACATATGC	77406
Qy	24	LeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyLeuThrGlyMet	43
Db	77407	CTCCGGGTTCAAGCCATCTCTCGCTCAACTCCCAAGTAGCTGGGATATACAGGTGTG	77466
Qy	44	CysThrHisAlaArgLeuLeuLeuThrPheLeuValGluMetGluPheLeuHisVal	63
Db	77467	CGCCACTATGCCCGCTAAIT---TTTATATTTTAGTAGATGAGGTTTCACCATGTT	77523
Qy	64	GlyGlnAlaGlyLeuGluLeuProThrSerAspProSerValSerAlaSerGlnSer	83
Db	77524	GGCCAGGCTTGTGCGAACTCTCTGAGCTGAGGTGATCCACCACCTCAGCCTCCCAAGT	77583
Qy	84	AlaArgTyrArgThr---GlyHisHisAlaArg-Leu	94
Db	77584	GCTGGGATTACAAGCATGAGCCACCGCGCGCGCTATATTTATTTATTTTATTTT	77643
Qy	94	-----	94
Db	77644	ACTTATTTATGACTTATTTATTTATTTATTTATTTATTTATTTATTTTGTGACACA	77703
Qy	94	-----	94
Db	77704	GTCTTGCTCTGTCACCAGGCTGAAGTGAATGGCATGATCTTGGCTCACTGCAACCCCC	77763
Qy	94	-----	94
Db	77764	ACTTCCAGGTTCAAGCAATCTCTGCTCAAGCTCCGAGTAGCTGGACTACAGGTG	77823
Qy	95	-----CysLeuAlaAsnPheCys-----GlyArgAsnArgValSerLeuMetCys	109
Db	77824	TGTGCCACCATGCTCGCTAAATTTTGTATTTTAGTAGATGGGTTTCA-----	77875
Qy	109	sProSerTrp-----SerProGluLeuIysGlnSerThrCysLeuSerLe	124
Db	77876	-CCATAGTGGCCAGGCTAAATCTCGAACTCTGAGCTCAGGTGATCGGCCACCTGGGCT	77934
Qy	124	uProIysCysTrpAspTyrArgArgAlaIalVal---ProGlyLeuPheIleLeuPhePh	143
Db	77935	CCCAAAATGCTGAGATTACAGGCGTGAGCCACACCCGTCCTCTTTTGTGT---TTTT	77991
Qy	143	eLeu-ArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAspHisSerS	163
Db	77992	TTTGGAGATGGGATCTCACTCTTTCCCAAGCCAGATACAGTGGAGAAATCATAGCT	78051
Qy	163	erLeuGlnProSerThrProGluIleIys-HisProProAlaSerAlaSerGlnValAla	182
Db	78052	CACTGCAGCATCAAACTCTCTGAGCTCAAGGGATCTCTCCAGCTCAGCCACCTAGTAGCT	78111
Qy	183	GlyThrIysAspMetHisIstYrThrTrpLeuIlePheIlePheAsnPheLeu	202
Db	78112	GGGATCAAGAGGTGTGACCAACCATGCCAGCTAAATTTGTTTAA---TTTTTTTTTTT	78170
Qy	203	ArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGlySerLeu	222
Db	78171	AGATGGAGTCTT---GCTGTACCCAGCTCGAGTGCAGTGTGCTCAGTCTCAGTCTAGT	78227
Qy	223	GlnProLeuProGlyPheIysLeuPheSerCysProSerLeuLeuSerSerTrpAsp	242

Search completed: September 15, 2005, 22:34:18  
Job time : 792 secs

[illegible]

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2005, 19:31:42 ; Search time 164 Seconds  
(without alignments)  
901.901 Million cell updates/sec

Title: US-09-380-203-2

Perfect score: 2034

Sequence: 1 MBFSLLPRLCNGAISARH.....FIRGGVSPYLQSGWQTPDLR 375

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1777461 seqs, 394431504 residues

Total number of hits satisfying chosen parameters: 1777461

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2034	100.0	375	9 US-09-964-666-2	Sequence 2, Appli
2	2034	100.0	375	9 US-09-964-412-2	Sequence 2, Appli
3	2034	100.0	375	10 US-09-964-667-2	Sequence 2, Appli
4	2034	100.0	375	10 US-09-872-968-2	Sequence 2, Appli
5	2034	100.0	375	10 US-09-964-678A-2	Sequence 2, Appli
6	2034	100.0	375	14 US-10-146-130-2	Sequence 2, Appli
7	2034	100.0	375	14 US-10-092-934-10	Sequence 10, Appli
8	2034	100.0	375	14 US-10-153-334-1	Sequence 1, Appli
9	2034	100.0	375	14 US-10-198-069-1	Sequence 1, Appli
10	2034	100.0	375	14 US-10-157-031-299	Sequence 299, App
11	2034	100.0	375	14 US-10-198-070-1	Sequence 1, Appli

12	2034	100.0	375	16 US-10-755-889-410	Sequence 410, App
13	2034	100.0	375	17 US-10-910-173-2	Sequence 2, Appli
14	734.5	36.1	361	9 US-09-995-494-107	Sequence 107, App
15	729	35.8	449	14 US-10-007-280A-140	Sequence 140, App
16	569	28.0	241	15 US-10-276-774-1834	Sequence 1834, Ap
17	517	25.4	213	15 US-10-296-115-911	Sequence 911, App
18	471	23.2	286	15 US-10-291-172-654	Sequence 654, App
19	471	23.2	286	15 US-10-221-278-654	Sequence 654, App
20	428.5	21.1	183	9 US-09-989-920-245	Sequence 245, App
21	407	20.0	361	15 US-10-276-774-1862	Sequence 1862, Ap
22	382.5	18.8	341	18 US-10-220-335-586	Sequence 586, App
23	382.5	18.8	673	14 US-10-157-031-291	Sequence 291, App
24	377.5	18.6	158	15 US-10-104-047-3047	Sequence 3047, Ap
25	356.5	17.5	603	16 US-10-408-765A-140	Sequence 140, App
26	353.5	17.4	217	14 US-10-017-161-1956	Sequence 1956, Ap
27	353.5	17.4	217	15 US-10-292-798-1604	Sequence 1604, Ap
28	349	17.2	157	15 US-10-108-260A-4272	Sequence 4272, Ap
29	336	16.5	148	15 US-10-276-774-1784	Sequence 1784, Ap
30	332	16.3	108	14 US-10-078-090-143	Sequence 143, App
31	324	15.9	189	15 US-10-104-047-3196	Sequence 3196, Ap
32	316	15.5	152	14 US-10-198-070-112	Sequence 112, App
33	311	15.3	74	15 US-10-276-774-1929	Sequence 1929, Ap
34	310	15.2	101	15 US-10-094-749-181	Sequence 2181, Ap
35	309.5	15.2	107	15 US-10-276-774-1874	Sequence 1874, Ap
36	308	15.1	93	10 US-09-764-891-5337	Sequence 5337, Ap
37	308	15.1	93	14 US-10-205-428-486	Sequence 486, App
38	308	15.1	100	13 US-10-016-157A-187	Sequence 187, App
39	307	15.1	114	15 US-10-104-047-2423	Sequence 2423, Ap
40	305.5	15.0	137	15 US-10-276-774-1959	Sequence 1959, Ap
41	303	14.9	203	15 US-10-108-260A-2748	Sequence 2748, Ap
42	302.5	14.9	128	15 US-10-276-774-2194	Sequence 2194, Ap
43	301	14.8	156	15 US-10-104-047-2334	Sequence 2334, Ap
44	300.5	14.8	101	15 US-10-276-774-2026	Sequence 2026, Ap
45	300	14.7	76	15 US-10-276-774-2032	Sequence 2032, Ap

#### ALIGNMENTS

#### RESULT 1

US-09-964-666-2

; Sequence 2, Application US/09964666

; Patent No. US20020104108A1

; GENERAL INFORMATION:

; APPLICANT: de la Monte, Suzanne

; TITLE OF INVENTION: Transgenic Animals and Cell Lines for

; Screening Drugs Effective for the Treatment or Prevention

; of Alzheimer's Disease

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.

; STREET: 1100 New York Ave., Suite 600

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20005-3934

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/964,666

; FILING DATE: 28-Sep-2001

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Esmond, Robert W.

; REGISTRATION NUMBER: 32,893

; REFERENCE/DOCKET NUMBER: 0609.4370000

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-371-2600

; TELEFAX: 202-371-2540

```
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 375 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-964-666-2

Query Match      100.0%; Score 2034; DB 9; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
DB 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
QY 61 LHVGQAGLELPTSDPSVASQASARYRTGHARLCLANFCGRNRSVLMCPSPSPKQST 120
DB 61 LHVGQAGLELPTSDPSVASQASARYRTGHARLCLANFCGRNRSVLMCPSPSPKQST 120
QY 121 CLSLPKCDYRRAAVPGLFILFRLHRCPRTLQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
DB 121 CLSLPKCDYRRAAVPGLFILFRLHRCPRTLQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
QY 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240
DB 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240
QY 241 WDYRRPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300
DB 241 WDYRRPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300
QY 301 LFEMESHVSTQAGVQWPNLGSLOPLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
DB 301 LFEMESHVSTQAGVQWPNLGSLOPLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
QY 361 VSPYLSGWSQTPDLR 375
DB 361 VSPYLSGWSQTPDLR 375

RESULT 2
US-09-964-412-2
; Sequence 2, Application US/09964412
; Patent No. US20020129391A1
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; Wanda, Jack R.
; TITLE OF INVENTION: Transgenic Animals and Cell Lines for
; Screening Drugs Effective for the Treatment or Prevention
; of Alzheimer's Disease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/964,412
; FILING DATE: 28-Sep-2001
; CLASSIFICATION: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0609.4370000
```

```
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 375 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-964-412-2

Query Match      100.0%; Score 2034; DB 9; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
DB 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
QY 61 LHVGQAGLELPTSDPSVASQASARYRTGHARLCLANFCGRNRSVLMCPSPSPKQST 120
DB 61 LHVGQAGLELPTSDPSVASQASARYRTGHARLCLANFCGRNRSVLMCPSPSPKQST 120
QY 121 CLSLPKCDYRRAAVPGLFILFRLHRCPRTLQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
DB 121 CLSLPKCDYRRAAVPGLFILFRLHRCPRTLQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
QY 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240
DB 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240
QY 241 WDYRRPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300
DB 241 WDYRRPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300
QY 301 LFEMESHVSTQAGVQWPNLGSLOPLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
DB 301 LFEMESHVSTQAGVQWPNLGSLOPLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
QY 361 VSPYLSGWSQTPDLR 375
DB 361 VSPYLSGWSQTPDLR 375

RESULT 3
US-09-964-667-2
; Sequence 2, Application US/09964667
; Publication No. US20030033621A1
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; Wanda, Jack R.
; TITLE OF INVENTION: Transgenic Animals and Cell Lines for
; Screening Drugs Effective for the Treatment or Prevention
; of Alzheimer's Disease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
; STREET: 1100 New York Ave., Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/964,667
; FILING DATE: 28-Sep-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
```

NAME: Esmond, Robert W.  
 REGISTRATION NUMBER: 32,893  
 REFERENCE/DOCKET NUMBER: 0609.4370000  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-371-2600  
 TELEFAX: 202-371-2540  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 375 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-09-964-667-2

Query Match 100.0%; Score 2034; DB 10; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 1e-184;  
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLCNGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60  
 DB 1 MEFSLLLPRLCNGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60

QY 61 LHVGQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRRVSLMCPSPSPKQST 120  
 DB 61 LHVGQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRRVSLMCPSPSPKQST 120

QY 121 CLSLPKCWDYRAAIVPGLFLLFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180  
 DB 121 CLSLPKCWDYRAAIVPGLFLLFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180

QY 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPPGKLFSCPSLLSS 240  
 DB 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPPGKLFSCPSLLSS 240

QY 241 WYRRPPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHARLIFNFC 300  
 DB 241 WYRRPPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHARLIFNFC 300

QY 301 LFEMESHVSVTQAGVQWRNLSLQPLPPGKLFSCPSLLSS 360  
 DB 301 LFEMESHVSVTQAGVQWRNLSLQPLPPGKLFSCPSLLSS 360

QY 361 VSPYLSGWSQTPDLR 375  
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 4  
 US-09-872-968-2  
 ; Sequence 2, Application US/09872968  
 ; Publication No. US20030050262A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wands, Jack R.  
 ; TITLE OF INVENTION: Inhibition of Neurodegeneration  
 ; FILE REFERENCE: 21496-047  
 ; CURRENT APPLICATION NUMBER: US/09/872,968  
 ; CURRENT FILING DATE: 2001-06-01  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 375  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; OTHER INFORMATION: AD7C-NTP CDNA  
 ; US-09-872-968-2

Query Match 100.0%; Score 2034; DB 10; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 1e-184;  
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLCNGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60  
 DB 1 MEFSLLLPRLCNGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60

QY 61 LHVGQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRRVSLMCPSPSPKQST 120  
 DB 61 LHVGQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRRVSLMCPSPSPKQST 120

QY 121 CLSLPKCWDYRAAIVPGLFLLFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180  
 DB 121 CLSLPKCWDYRAAIVPGLFLLFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180

QY 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPPGKLFSCPSLLSS 240  
 DB 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPPGKLFSCPSLLSS 240

QY 241 WYRRPPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHARLIFNFC 300  
 DB 241 WYRRPPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHARLIFNFC 300

QY 301 LFEMESHVSVTQAGVQWRNLSLQPLPPGKLFSCPSLLSS 360  
 DB 301 LFEMESHVSVTQAGVQWRNLSLQPLPPGKLFSCPSLLSS 360

QY 361 VSPYLSGWSQTPDLR 375  
 DB 361 VSPYLSGWSQTPDLR 375

QY 1 MEFSLLLPRLCNGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60  
 DB 1 MEFSLLLPRLCNGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60

QY 61 LHVGQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRRVSLMCPSPSPKQST 120  
 DB 61 LHVGQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRRVSLMCPSPSPKQST 120

QY 121 CLSLPKCWDYRAAIVPGLFLLFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180  
 DB 121 CLSLPKCWDYRAAIVPGLFLLFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180

QY 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPPGKLFSCPSLLSS 240  
 DB 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPPGKLFSCPSLLSS 240

QY 241 WYRRPPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHARLIFNFC 300  
 DB 241 WYRRPPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHARLIFNFC 300

QY 301 LFEMESHVSVTQAGVQWRNLSLQPLPPGKLFSCPSLLSS 360  
 DB 301 LFEMESHVSVTQAGVQWRNLSLQPLPPGKLFSCPSLLSS 360

QY 361 VSPYLSGWSQTPDLR 375  
 DB 361 VSPYLSGWSQTPDLR 375

RESULT 5  
 US-09-964-678A-2  
 ; Sequence 2, Application US/09964678A  
 ; Publication No. US2003006097A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wands, Jack R.  
 ; TITLE OF INVENTION: Transgenic Animals and Cell Lines for Screening Drugs  
 ; TITLE OF INVENTION: Effective for the Treatment or Prevention of  
 ; TITLE OF INVENTION: Alzheimer's Disease  
 ; FILE REFERENCE: 0609.4370002  
 ; CURRENT APPLICATION NUMBER: US/09/964,678A  
 ; CURRENT FILING DATE: 2001-09-28  
 ; PRIOR APPLICATION NUMBER: 09/380,203  
 ; PRIOR FILING DATE: 2000-04-25  
 ; PRIOR APPLICATION NUMBER: PCT/US98/03685  
 ; PRIOR FILING DATE: 1998-02-26  
 ; PRIOR APPLICATION NUMBER: 60/038,908  
 ; PRIOR FILING DATE: 1997-02-26  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: Patent In version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 375  
 ; TYPE: PRT  
 ; ORGANISM: Unknown  
 ; FEATURE:  
 ; OTHER INFORMATION: AD7C-NTP CDNA  
 ; US-09-964-678A-2

Query Match 100.0%; Score 2034; DB 10; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 1e-184;  
 Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPRLCNGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60  
 DB 1 MEFSLLLPRLCNGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60

QY 61 LHVGQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRRVSLMCPSPSPKQST 120  
 DB 61 LHVGQAGLELPTSDPSVSASQARYRTGHARLCLANFCGRRVSLMCPSPSPKQST 120

QY 121 CLSLPKCWDYRAAIVPGLFLLFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180  
 DB 121 CLSLPKCWDYRAAIVPGLFLLFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPASASQ 180

QY 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPPGKLFSCPSLLSS 240  
 DB 181 VAGTKDMHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPPGKLFSCPSLLSS 240

```
Db 181 VAGTKDMHHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240
Qy 241 WDYRRPPRLANFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300
Db 241 WDYRRPPRLANFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300
Qy 301 LFEMESHVSTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 360
Db 301 LFEMESHVSTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 360
Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375
```

## RESULT 6

```
US-10-146-130-2
; Sequence 2, Application US/10146130
; Publication No. US20030004107A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: METHOD OF PREVENTING CELL DEATH USING SEGMENTS OF
; FILE REFERENCE: 59003.00007
; CURRENT APPLICATION NUMBER: US/10/146,130
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-130-2
```

```
Query Match 100.0%; Score 2034; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Qy 61 LHVGQAGLELPTSDPSVSASQARYTGHARLCLANFCGRRNVSMLCPSWSPKQST 120
Db 61 LHVGQAGLELPTSDPSVSASQARYTGHARLCLANFCGRRNVSMLCPSWSPKQST 120
Qy 121 CLSLPKCDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
Db 121 CLSLPKCDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
Qy 181 VAGTKDMHHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240
Db 181 VAGTKDMHHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240
Qy 241 WDYRRPPRLANFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300
Db 241 WDYRRPPRLANFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300
Qy 301 LFEMESHVSTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 360
Db 301 LFEMESHVSTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 360
Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375
```

```
US-10-153-334-1
; Sequence 1, Application US/10153334
; Publication No. US20030096350A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; FILE REFERENCE: 59003-00006
; CURRENT APPLICATION NUMBER: US/10/153,334
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-334-1
```

## RESULT 7

```
US-10-092-934-10
; Sequence 10, Application US/10092934
; Publication No. US20030054990A1
; GENERAL INFORMATION:
```

```
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: METHODS OF USING NEURAL THREAD PROTEINS TO TREAT TUMORS
; FILE REFERENCE: 018792-0199
; CURRENT APPLICATION NUMBER: US/10/092,934
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/273,957
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-934-10
```

```
Query Match 100.0%; Score 2034; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Db 1 MEFSLLLPRLECNCGAISAHNRNLRPGSSDSPASASPVAGITGCTHARLILYFFLVEMEF 60
Qy 61 LHVGQAGLELPTSDPSVSASQARYTGHARLCLANFCGRRNVSMLCPSWSPKQST 120
Db 61 LHVGQAGLELPTSDPSVSASQARYTGHARLCLANFCGRRNVSMLCPSWSPKQST 120
Qy 121 CLSLPKCDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
Db 121 CLSLPKCDYRRAAVPGLFILFRLHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180
Qy 181 VAGTKDMHHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240
Db 181 VAGTKDMHHYTWLIFIFNFRQSLNSVTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 240
Qy 241 WDYRRPPRLANFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300
Db 241 WDYRRPPRLANFVFLVEMGFTMFARLILISGCDLPASASQAGITGVSHHARLIFNFC 300
Qy 301 LFEMESHVSTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 360
Db 301 LFEMESHVSTQAGVQWRNLGSLQPLPPGKLFSCPSLLSS 360
Qy 361 VSPYLSGWSQTPDLR 375
Db 361 VSPYLSGWSQTPDLR 375
```

## RESULT 8

```
US-10-153-334-1
; Sequence 1, Application US/10153334
; Publication No. US20030096350A1
; GENERAL INFORMATION:
```

```
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; FILE REFERENCE: 59003-00006
; CURRENT APPLICATION NUMBER: US/10/153,334
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-334-1
```

```
Query Match 100.0%; Score 2034; DB 14; Length 375;
```

Best Local Similarity 100.0%; Pred. No. 1e-184;  
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLPRLCNGAISAHNRNLRPGSSDPSASPVAGITGCMCTHARLILYFFLVEMEF 60  
DB 1 MEFSLLPRLCNGAISAHNRNLRPGSSDPSASPVAGITGCMCTHARLILYFFLVEMEF 60

QY 61 LHVGAGLELPTSDPSVSASQARYRTGHHARLCLANFCGRNRYSLMCPSPKOST 120  
DB 61 LHVGAGLELPTSDPSVSASQARYRTGHHARLCLANFCGRNRYSLMCPSPKOST 120

QY 121 CLSLPKWDYRRAAVPGLFILFLLHRCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180  
DB 121 CLSLPKWDYRRAAVPGLFILFLLHRCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180

QY 181 VAGTKDMHHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPPGFKLSPSCSLSS 240  
DB 181 VAGTKDMHHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPPGFKLSPSCSLSS 240

QY 241 WDYRRPRLANFFVFLVEMGTFMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300  
DB 241 WDYRRPRLANFFVFLVEMGTFMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300

QY 301 LFEMESHVTOAGVQWPNLGSLOPLPPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGG 360  
DB 301 LFEMESHVTOAGVQWPNLGSLOPLPPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGG 360

QY 361 VSPYLSGWSQTPDLR 375  
DB 361 VSPYLSGWSQTPDLR 375

RESULT 9  
US-10-198-069-1  
; Sequence 1, Application US/10198069  
; Publication No. US20030096756A1  
; GENERAL INFORMATION:  
; APPLICANT: AVERBACK, PAUL  
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER  
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF  
; TITLE OF INVENTION: CELLS  
; FILE REFERENCE: 59003.000009  
; CURRENT APPLICATION NUMBER: US/10/198,069  
; CURRENT FILING DATE: 2002-07-19  
; PRIOR APPLICATION NUMBER: 60/306,161  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: 60/306,150  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: 60/331,477  
; PRIOR FILING DATE: 2001-11-16  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-198-069-1

Query Match 100.0%; Score 2034; DB 14; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1e-184;  
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLPRLCNGAISAHNRNLRPGSSDPSASPVAGITGCMCTHARLILYFFLVEMEF 60  
DB 1 MEFSLLPRLCNGAISAHNRNLRPGSSDPSASPVAGITGCMCTHARLILYFFLVEMEF 60

QY 61 LHVGAGLELPTSDPSVSASQARYRTGHHARLCLANFCGRNRYSLMCPSPKOST 120  
DB 61 LHVGAGLELPTSDPSVSASQARYRTGHHARLCLANFCGRNRYSLMCPSPKOST 120

QY 121 CLSLPKWDYRRAAVPGLFILFLLHRCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180  
DB 121 CLSLPKWDYRRAAVPGLFILFLLHRCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180

QY 181 VAGTKDMHHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPPGFKLSPSCSLSS 240  
DB 181 VAGTKDMHHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPPGFKLSPSCSLSS 240

QY 241 WDYRRPRLANFFVFLVEMGTFMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300  
DB 241 WDYRRPRLANFFVFLVEMGTFMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300

QY 301 LFEMESHVTOAGVQWPNLGSLOPLPPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGG 360  
DB 301 LFEMESHVTOAGVQWPNLGSLOPLPPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGG 360

QY 361 VSPYLSGWSQTPDLR 375  
DB 361 VSPYLSGWSQTPDLR 375

RESULT 10  
US-10-157-031-299  
; Sequence 299, Application US/10157031  
; Publication No. US20030108890A1  
; GENERAL INFORMATION:  
; APPLICANT: Baranova, A. V.  
; APPLICANT: Yankovsky, N. K.  
; APPLICANT: Kozlov, A. P.  
; APPLICANT: Lobashev, A. V.  
; APPLICANT: Krukovskaya, L. L.  
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences  
; FILE REFERENCE: 2760-103  
; CURRENT APPLICATION NUMBER: US/10/157,031  
; CURRENT FILING DATE: 2002-05-30  
; NUMBER OF SEQ ID NOS: 415  
; SOFTWARE: Patent In version 3.1  
; SEQ ID NO 299  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-157-031-299

Query Match 100.0%; Score 2034; DB 14; Length 375;  
Best Local Similarity 100.0%; Pred. No. 1e-184;  
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLPRLCNGAISAHNRNLRPGSSDPSASPVAGITGCMCTHARLILYFFLVEMEF 60  
DB 1 MEFSLLPRLCNGAISAHNRNLRPGSSDPSASPVAGITGCMCTHARLILYFFLVEMEF 60

QY 61 LHVGAGLELPTSDPSVSASQARYRTGHHARLCLANFCGRNRYSLMCPSPKOST 120  
DB 61 LHVGAGLELPTSDPSVSASQARYRTGHHARLCLANFCGRNRYSLMCPSPKOST 120

QY 121 CLSLPKWDYRRAAVPGLFILFLLHRCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180  
DB 121 CLSLPKWDYRRAAVPGLFILFLLHRCPTLTQDEVQWCHSSLOPSTPEIKHPPASASQ 180

QY 181 VAGTKDMHHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPPGFKLSPSCSLSS 240  
DB 181 VAGTKDMHHYTWLIFIFNFRQSLNSVTQAGVQWRNLSLQPLPPGFKLSPSCSLSS 240

QY 241 WDYRRPRLANFFVFLVEMGTFMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300  
DB 241 WDYRRPRLANFFVFLVEMGTFMFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300

QY 301 LFEMESHVTOAGVQWPNLGSLOPLPPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGG 360  
DB 301 LFEMESHVTOAGVQWPNLGSLOPLPPGLKRFSCSLSPSSWDYGHLPHPANFCIFIRGG 360

QY 361 VSPYLSGWSQTPDLR 375  
DB 361 VSPYLSGWSQTPDLR 375



```
RESULT 11
US-10-198-070-1
; Sequence 1, Application US/10198070
; Publication No. US20030109437A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; APPLICANT: GEMMELL, JACK
; TITLE OF INVENTION: PETIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 59003.000008
; CURRENT APPLICATION NUMBER: US/10/198,070
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 60/306,161
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/306,150
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/331,477
; PRIOR FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-070-1

Query Match      100.0%; Score 2034; DB 14; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPLECNCGAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFFLVEMEF 60
DB 1 MEFSLLLPLECNCGAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFFLVEMEF 60
QY 61 LHVGQAGLELPTSDPSVSASOSARYRTGHARLCLANFCGRRNVRSLMCPSPSPKQST 120
DB 61 LHVGQAGLELPTSDPSVSASOSARYRTGHARLCLANFCGRRNVRSLMCPSPSPKQST 120
QY 121 CLSLPKCDYRRAAVPGLFIFLFRHRCPTLTQDEVQVQWCHSSLOPSTPEIKHPPASASQ 180
DB 121 CLSLPKCDYRRAAVPGLFIFLFRHRCPTLTQDEVQVQWCHSSLOPSTPEIKHPPASASQ 180
QY 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWNLGSLQPLPPGKLFSCPSLLSS 240
DB 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWNLGSLQPLPPGKLFSCPSLLSS 240
QY 241 WDYRPPRLANFFVFLVEMGFTMFARLILISGPCDLPASASOSAGITGVSHHARLIFNFC 300
DB 241 WDYRPPRLANFFVFLVEMGFTMFARLILISGPCDLPASASOSAGITGVSHHARLIFNFC 300
QY 301 LFEMESHVSTQAGVQWPNLGSLOPLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
DB 301 LFEMESHVSTQAGVQWPNLGSLOPLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
QY 361 VSPYLSGWSQTPDLR 375
DB 361 VSPYLSGWSQTPDLR 375

RESULT 12
US-10-755-889-410
; Sequence 410, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14

; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 410
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-410

Query Match      100.0%; Score 2034; DB 16; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPLECNCGAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFFLVEMEF 60
DB 1 MEFSLLLPLECNCGAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFFLVEMEF 60
QY 61 LHVGQAGLELPTSDPSVSASOSARYRTGHARLCLANFCGRRNVRSLMCPSPSPKQST 120
DB 61 LHVGQAGLELPTSDPSVSASOSARYRTGHARLCLANFCGRRNVRSLMCPSPSPKQST 120
QY 121 CLSLPKCDYRRAAVPGLFIFLFRHRCPTLTQDEVQVQWCHSSLOPSTPEIKHPPASASQ 180
DB 121 CLSLPKCDYRRAAVPGLFIFLFRHRCPTLTQDEVQVQWCHSSLOPSTPEIKHPPASASQ 180
QY 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWNLGSLQPLPPGKLFSCPSLLSS 240
DB 181 VAGTKDMHYTWLIFIFINFLRQSLNSVTQAGVQWNLGSLQPLPPGKLFSCPSLLSS 240
QY 241 WDYRPPRLANFFVFLVEMGFTMFARLILISGPCDLPASASOSAGITGVSHHARLIFNFC 300
DB 241 WDYRPPRLANFFVFLVEMGFTMFARLILISGPCDLPASASOSAGITGVSHHARLIFNFC 300
QY 301 LFEMESHVSTQAGVQWPNLGSLOPLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
DB 301 LFEMESHVSTQAGVQWPNLGSLOPLPPGLKRFSCLSLPSSWDYGHLPHPANFCIFIRGG 360
QY 361 VSPYLSGWSQTPDLR 375
DB 361 VSPYLSGWSQTPDLR 375

; QUERY MATCH
; BEST LOCAL SIMILARITY
; MATCHES
; CONSERVATIVE
; MISMATCHES
; INDELS
; GAPS
; SEQUENCE 1, APPLICATION US/10910173
; PUBLICATION NO. US20050090441A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R
; APPLICANT: de la Monte, Suzanne M
; TITLE OF INVENTION: Inhibition of Neurodegeneration
; FILE REFERENCE: 21486-047
; CURRENT APPLICATION NUMBER: US/10/910,173
; CURRENT FILING DATE: 2004-08-02
; PRIOR APPLICATION NUMBER: US/09/872,968
; PRIOR FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-910-173-2

Query Match      100.0%; Score 2034; DB 17; Length 375;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFSLLLPLECNCGAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFFLVEMEF 60
DB 1 MEFSLLLPLECNCGAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFFLVEMEF 60
QY 61 LHVGQAGLELPTSDPSVSASOSARYRTGHARLCLANFCGRRNVRSLMCPSPSPKQST 120
```

Db 61 LRVQAGLEPTSDPSVSASQARYRTGHHARLCLANFCGRNRVLMCPSPWSPKQST 120  
QY 121 CLSLPKCWDYRRAAVPGLFLLPFLRHRCPTLTODEVQWCDHSLQSTPEIKHPPASQ 180  
Db 121 CLSLPKCWDYRRAAVPGLFLLPFLRHRCPTLTODEVQWCDHSLQSTPEIKHPPASQ 180  
QY 181 VAGTKDMHYTWMIFIFINFLRQSLNSVTQAGVQWRNLGSLQPLPPGFKLFCSPSLSS 240  
Db 181 VAGTKDMHYTWMIFIFINFLRQSLNSVTQAGVQWRNLGSLQPLPPGFKLFCSPSLSS 240  
QY 241 WDYRRPRLANPFVFLVEMGFTWFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300  
Db 241 WDYRRPRLANPFVFLVEMGFTWFARLILISGPCDLPASASQAGITGVSHHARLIFNFC 300  
QY 301 LFEMESHVSTQAGVQWRNLGSLQPLPPGFKLFCSPSLSS 360  
Db 301 LFEMESHVSTQAGVQWRNLGSLQPLPPGFKLFCSPSLSS 360  
QY 361 VSPYLSGWSQTPDLR 375  
Db 361 VSPYLSGWSQTPDLR 375

## RESULT 14

US-09-995-494-107  
; Sequence 107, Application US/09995494  
; Patent No. US20020127578A1  
; GENERAL INFORMATION:  
; APPLICANT: Salceda, Susana  
; APPLICANT: Macina, Roberto  
; APPLICANT: Recipon, Herve  
; APPLICANT: Cafferkey, Robert  
; APPLICANT: Ali, Shujath  
; APPLICANT: Sun, Yongming  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Chen, Sei-Yu  
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and  
; FILE REFERENCE: DEX-0293  
; CURRENT APPLICATION NUMBER: US/09/995,494  
; CURRENT FILING DATE: 2001-11-27  
; PRIOR APPLICATION NUMBER: 60/253,176  
; PRIOR FILING DATE: 2000-11-27  
; NUMBER OF SEQ ID NOS: 115  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 107  
; LENGTH: 361  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-995-494-107

Query Match 36.1%; Score 734.5; DB 9; Length 361;  
Best Local Similarity 43.1%; Pred. No. 4.3e-61;  
Matches 185; Conservative 24; Mismatches 93; Indels 127; Gaps 12;  
QY 6 LLPRLECNAGISAHRNLRLPGSSDSPASASPVAGITGM-----CTHARL----- 49  
Db 1 LLPRLECSGTIAHCHLKLGGDLPLASASRVGGITGMQPTMGTSCHAFQIGLPPVAN 60  
QY 50 -----ILYFVLEMEFLHVQAGLELPTSDPSVSASQARYRTGHHARLCL 97  
Db 61 PERSFYLGTSPLDOLWLNMYF-----FALLIHIVLFL- 94  
QY 98 NFGCNRVLMCP--SWSPELQSTCLSLPKCWDYRRAAVPGLFLLFHLHRC----- 148  
Db 95 -----NRDSCCPGASLTGLGFAFACLSLPKCWDYTYGELL-LPIDIFLQVCLSFYFLN 148  
QY 149 PTLTODEVQWCDHSLQSTPEIKHPPASASQVAGTKDMHYTWMIFIFINFLRQSLNS 208  
Db 149 TVFORAEV-----LIFFQWVYLAWSLHS 172  
QY 209 VTOAGVQWRNLGSLQPLPPGFKLFCSPSLSSWDYRR--PPRLAN--FFVFLVEMGFTMPA 265

Db 173 VAQAGVQWCMNLGSLQPLPPRFRFSCLSLSSWDHRRHAPPCLANFLFFKFLVDQSFMTLA 232  
QY 266 RLILISGPC-DLPASASQAGITGVSH-----HARLIFNFCLFEMES 306  
Db 233 RUVLNSASGDDLPAPASQAGITGVHCTWPKSKFKFADSHIGLAFHFAFFFFFAVAS 292  
QY 307 HSVTQAGVQWRNLGSLQPLPPGFKLFCSPSLSSWDYGHLPHPHPANFCIFIRGGVSPYLS 366  
Db 293 HFIAQAGVQWRDLGSLQPLPPGFKQFLCLSLPGSDYRRAPPRQANFCIFSRDGVSPCWT 352  
QY 367 GWSQTPDLR 375  
Db 353 GWSQTPDLR 361

## RESULT 15

US-10-007-280A-140  
; Sequence 140, Application US/10007280A  
; Publication No. US20030059784A1  
; GENERAL INFORMATION:  
; APPLICANT: Sun, Yongming  
; APPLICANT: Recipon, Herve  
; APPLICANT: Salceda, Susana  
; APPLICANT: Chenghua, Liu  
; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and  
; FILE REFERENCE: DEX-0257  
; CURRENT APPLICATION NUMBER: US/10/007,280A  
; CURRENT FILING DATE: 2001-11-07  
; PRIOR APPLICATION NUMBER: US 60/246,640  
; PRIOR FILING DATE: 2000-11-08  
; NUMBER OF SEQ ID NOS: 238  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 140  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-007-280A-140

Query Match 35.8%; Score 729; DB 14; Length 449;  
Best Local Similarity 42.6%; Pred. No. 1.9e-60;  
Matches 196; Conservative 40; Mismatches 122; Indels 102; Gaps 19;  
QY 6 LLPRLECNAGISAHRNLRLPGSSDSPASASPVAGITGMCTHARL-----ILYFVLE 57  
Db 1 LPFRLEYGTTILAYCNLHLPSSNPPTSASQVAGTRDVCHHTLWLCVVCVCVCVCVCVE 60  
QY 58 MEFLHVQAGLELPTSDPSVSASQARYRTGHHARLC-----LANFC----- 100  
Db 61 MRHYVQAGLELSSDDPPIASASQAGIIGISH---CTWPMHDSFISPGAEPLTFAYTW 117  
QY 101 -GRNRVSL-----MCPSPWSPKQSTCLSLPKCWDYRR-----AAVPGI-----F 139  
Db 118 PGRPKIPITILLYPGDDVLVAFRTELYASPSQPGASDTARESWNGAVPDLHKEW 177  
QY 140 ILFFLHRHCPTLTQDEV-----QWCD--HSSLQSTPEIKHPPA-----SASQVAGTKD 186  
Db 178 LIF-----CPFSNQSHLWTKSKWAEVPHPGRAELPAKKEQKAANENSGSVTEPSSAS 232  
QY 187 MHYTWLFIIFN-----FLRQSLNSVTQAGVQW---RNLSLQPLPPGFKLFCSPSLSS 239  
Db 233 ILHARWDYV-FLINALIYFLRQSLRSVAQAGVQWCGADLQSLQPLPPGKAFPCLLSL 291  
QY 240 .SWDYRR--PPRLANPFVFLVEMGFTWFARLILISGPCDLPASASQAGITGVSHHAR---- 294  
Db 292 SWDYRSLEPCPANPFVFLIETGFHHISQ-ISISAFCDPPASASQAGITGSHCAQPDVY 350  
QY 295 -----LIFNFCLFEMESHVSTQAGVQWPNLG-----SLOPLPPGLKRFSC 334  
Db 351 YVSVGIQKQRCYLFVFFFFETESRTVAQAG-RLERSGAISTRRLQPLPPGLKRFSC 409  
QY 335 LSLSSWDYGHLPHPHPANFCIFIRGGVSPYLSGWSQTPDL 374  
Db 410 LSLSSWDYRCTPPRLAHFCTFSRDGVSPCWGWSLSPDL 449

Search completed: September 13, 2005, 19:38:05  
Job time : 166 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2005, 19:32:27 ; Search time 43 Seconds  
(without alignments)  
651.009 Million cell updates/sec

Title: US-09-380-203-2  
Perfect score: 2034  
Sequence: 1 MEFSLLPRLCNGAISAHRLPGSSDSDSPASPVAGITGMCTHARLILYFFLVAMEF 60

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2034	100.0	375	2	US-08-454-557C-121
2	2034	100.0	375	2	US-08-340-426D-121
3	2034	100.0	375	2	US-08-450-673C-121
4	2034	100.0	375	4	US-09-872-968-2
5	1415.5	69.6	397	5	PCT-US95-17111A-121
6	267.5	13.2	96	4	US-09-513-999C-6065
7	259.5	12.8	144	4	US-09-513-999C-6953
8	256	12.6	108	4	US-09-513-999C-7878
9	254.5	12.5	239	4	US-09-800-729-193
10	254.5	12.5	310	4	US-09-800-729-190
11	253	12.4	118	4	US-09-663-600A-114
12	251	12.3	776	4	US-10-020-079-24
13	251	12.3	789	4	US-10-020-079-22
14	251	12.3	863	4	US-10-020-079-32
15	251	12.3	876	4	US-10-020-079-30
16	251	12.3	889	4	US-10-020-079-20
17	251	12.3	895	4	US-10-020-079-18
18	251	12.3	976	4	US-10-020-079-28
19	251	12.3	982	4	US-10-020-079-26
20	250	12.3	97	4	US-09-513-999C-4770
21	250	12.3	132	4	US-09-636-215-573
22	250	12.3	132	4	US-09-685-166A-573
23	250	12.3	132	4	US-09-679-426-573
24	250	12.3	132	4	US-09-759-143-573
25	250	12.3	132	4	US-09-651-236-573
26	250	12.3	135	4	US-09-685-166A-884
27	250	12.3	135	4	US-09-679-426-884

28	250	12.3	135	4	US-09-759-143-884	Sequence 884, App
29	247.5	12.2	112	4	US-09-513-999C-7870	Sequence 7870, App
30	246	12.1	119	4	US-09-513-999C-7867	Sequence 7867, App
31	245	12.0	102	4	US-09-621-976-6112	Sequence 6112, App
32	242.5	11.9	121	4	US-09-513-999C-7874	Sequence 7874, App
33	241	11.8	91	4	US-09-621-976-5929	Sequence 5929, App
34	229	11.3	84	4	US-09-621-976-7167	Sequence 7167, App
35	226	11.1	61	4	US-09-513-999C-4581	Sequence 4581, App
36	224.5	11.0	396	4	US-09-949-016-6783	Sequence 6783, App
37	224.5	11.0	442	4	US-09-949-016-9762	Sequence 9762, App
38	223.5	11.0	87	4	US-09-621-976-5968	Sequence 5968, App
39	219	10.8	76	4	US-09-621-976-6338	Sequence 6338, App
40	218	10.7	1079	3	US-09-058-489-22	Sequence 22, Appl
41	217.5	10.7	87	4	US-09-205-258-342	Sequence 342, App
42	216	10.6	169	4	US-09-663-600A-208	Sequence 208, App
43	215.5	10.6	108	4	US-09-513-999C-4433	Sequence 4433, App
44	214.5	10.5	83	4	US-09-621-976-5396	Sequence 5396, App
45	211.5	10.4	103	4	US-09-513-999C-5327	Sequence 5327, App

ALIGNMENTS

RESULT 1

US-08-454-557C-121  
; Sequence 121, Application US/08454557C  
; Patent No. 5830670  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection  
; TITLE OF INVENTION: Of Alzheimer's Disease  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/454,557C  
; FILING DATE: 30-MAY-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ludwig, Steven R.  
; REGISTRATION NUMBER: 36,203  
; REFERENCE/DOCKET NUMBER: 0809.3840003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 121:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 375 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-454-557C-121

Query Match 100.0%; Score 2034; DB 2; Length 375;  
Best Local Similarity 100.0%; Pred. No. 6.2e-215;  
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MEFSLLPRLCNGAISAHRLPGSSDSDSPASPVAGITGMCTHARLILYFFLVAMEF	60
Db	1	MEFSLLPRLCNGAISAHRLPGSSDSDSPASPVAGITGMCTHARLILYFFLVAMEF	60
Qy	61	LHVQAGLELTPSDPSVSASQSAARYRTGTHARLCLANFCGRNVRVSLMCPSPSPKQST	120

Db 61 LVHQAGLELPTSDPSVSASQARYTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120  
121 CLSLPKWDYRAAIVPGLFILFHLRHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180  
121 CLSLPKWDYRAAIVPGLFILFHLRHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180  
181 VAGTKDMHHYTWLIFIFIFNLRQSLNSVTQAGVQWRNLSLOPLPPGPKLFCSPSLSS 240  
181 VAGTKDMHHYTWLIFIFIFNLRQSLNSVTQAGVQWRNLSLOPLPPGPKLFCSPSLSS 240  
241 WDYRPPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIENPC 300  
241 WDYRPPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIENPC 300  
301 LFEMESHVSQTQAGVQWRNLSLOPLPPGPKLFCSPSLSS 360  
301 LFEMESHVSQTQAGVQWRNLSLOPLPPGPKLFCSPSLSS 360  
361 VSPYLSGWSQTPDLR 375  
361 VSPYLSGWSQTPDLR 375

RESULT 2  
US-08-340-426D-121  
; Sequence 121, Application US/08340426D  
; Patent No. 5948634  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; APPLICANT: Wands, Jack R.  
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection  
; TITLE OF INVENTION: of Alzheimer's Disease  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 14-NOV-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ludwig, Steven R.  
; REGISTRATION NUMBER: 36,203  
; REFERENCE/DOCKET NUMBER: 0609.3840002  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 121:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 375 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-340-426D-121

Query Match 100.0%; Score 2034; DB 2; Length 375;  
Best Local Similarity 100.0%; Pred. No. 6.2e-215;  
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEFSLLLPLECNGAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60  
Db 1 MEFSLLLPLECNGAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60

QY 61 LVHQAGLELPTSDPSVSASQARYTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120  
Db 61 LVHQAGLELPTSDPSVSASQARYTGHARLCLANFCGRNRVSLMCPSPWSPKQST 120  
QY 121 CLSLPKWDYRAAIVPGLFILFHLRHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180  
Db 121 CLSLPKWDYRAAIVPGLFILFHLRHRCPTLTQDEVQWCDHSSLOPSTPEIKHPPASASQ 180  
QY 181 VAGTKDMHHYTWLIFIFIFNLRQSLNSVTQAGVQWRNLSLOPLPPGPKLFCSPSLSS 240  
Db 181 VAGTKDMHHYTWLIFIFIFNLRQSLNSVTQAGVQWRNLSLOPLPPGPKLFCSPSLSS 240  
QY 241 WDYRPPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIENPC 300  
Db 241 WDYRPPRLANFVFLVEMGFTMFARLILISGPCDLPASASQAGITGVSHHARLIENPC 300  
QY 301 LFEMESHVSQTQAGVQWRNLSLOPLPPGPKLFCSPSLSS 360  
Db 301 LFEMESHVSQTQAGVQWRNLSLOPLPPGPKLFCSPSLSS 360  
QY 361 VSPYLSGWSQTPDLR 375  
Db 361 VSPYLSGWSQTPDLR 375

RESULT 3  
US-08-450-673C-121  
; Sequence 121, Application US/08450673C  
; Patent No. 5948888  
; GENERAL INFORMATION:  
; APPLICANT: de la Monte, Suzanne  
; APPLICANT: Wands, Jack R.  
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection  
; TITLE OF INVENTION: of Alzheimer's Disease  
; NUMBER OF SEQUENCES: 121  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/450,673C  
; FILING DATE: 30-MAY-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ludwig, Steven R.  
; REGISTRATION NUMBER: 36,203  
; REFERENCE/DOCKET NUMBER: 0609.3840004  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 121:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 375 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-450-673C-121

Query Match 100.0%; Score 2034; DB 2; Length 375;  
Best Local Similarity 100.0%; Pred. No. 6.2e-215;  
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MEFSLLLPLECNGAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60  
Db 1 MEFSLLLPLECNGAISAHNRNLRPGSSDSPASASPVAGITGMCTHARLILYFFLVEMEF 60



Db 282 --PKVLGLQDVPTARPINFCLEFEMESHVTTQAGVQWENLGLQLPPLPGLKRFSCLSLP 339  
QY 339 SSWDYGHLPHPANFCIFIRGG 360  
Db 340 SSWDYGHLPHTPLIFVFSLEAG 361

## RESULT 6

US-09-513-999C-6065  
; Sequence 6065, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961

; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6065  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-09-513-999C-6065

Query Match 13.2%; Score 267.5; DB 4; Length 96;  
Best Local Similarity 60.4%; Pred. No. 1.5e-21;  
Matches 61; Conservative 3; Mismatches 22; Indels 15; Gaps 2;

QY 205 SLNSVTQAGVQWRNLGSLQPLPGFKLFSCPSLLSSWDYRRPPRLANFFVFLVEMGFTWF 264  
Db 2 SFTLVAQAGVQWRDLGSPPLPGFKFSCPSLPSWDYRHPVLLPANLVFLVEMGF--- 58

QY 265 AKLLISGPC-----DLPASQSAGITGVSHARLIF 297  
Db 59 ----LHVGPAGLGLPTSGDLPASQSAGIIGVSHRSQPF 95

## RESULT 7

US-09-513-999C-6953  
; Sequence 6953, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961

; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6953  
; LENGTH: 144  
; TYPE: PRT  
; ORGANISM: Homo sapiens

; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 9  
; OTHER INFORMATION: Xaa= \* or Cys or Trp  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 10  
; OTHER INFORMATION: Xaa=Leu or Met or Val

; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 72  
; OTHER INFORMATION: Xaa=Glu or Val  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 110  
; OTHER INFORMATION: Xaa=Asp or His or Asn or Tyr  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 132  
; OTHER INFORMATION: Xaa=Asn or Thr  
US-09-513-999C-6953

Query Match 12.8%; Score 259.5; DB 4; Length 144;  
Best Local Similarity 72.0%; Pred. No. 2e-20;  
Matches 59; Conservative 3; Mismatches 19; Indels 1; Gaps 1;

QY 3 FSLILPRLCNGAISAHENLRIPGSSDSPASVPAGITGMCTHARLILYPFLVEMEPFLH 62  
Db 23 FSLLSRPRECNCGVILAHNCNLRLSGTHSPVSASGVAGITGMHHAQLI-FXFLVETGPHH 81

QY 63 VQAGLELPTSDDPVSASQSA 84  
Db 82 VQAGLELLTSGDPPASASQSA 103

## RESULT 8

US-09-513-999C-7878  
; Sequence 7878, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961

; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 7878  
; LENGTH: 108  
; TYPE: PRT  
; ORGANISM: Homo sapiens

; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -17..-1  
; OTHER INFORMATION: score 5.7  
; OTHER INFORMATION: seq VVCCSCSWFFLFC/FV  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 56  
; OTHER INFORMATION: Xaa= \* or Cys or Gly  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 59  
; OTHER INFORMATION: Xaa=Arg or Ser

; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 61  
; OTHER INFORMATION: Xaa=Asp or Asn  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 62  
; OTHER INFORMATION: Xaa=Ala or Glu or Lys or Met or Arg or Thr or Val  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 64  
; OTHER INFORMATION: Xaa= \* or Leu



FEATURE:  
NAME/KEY: UNSURE  
LOCATION: 66  
OTHER INFORMATION: Xaa=Asn or Tyr  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: 67  
OTHER INFORMATION: Xaa= \* or Tyr  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: 71  
OTHER INFORMATION: Xaa=Pro or Ser  
US-09-513-999C-7878

Query Match 12.6%; Score 256; DB 4; Length 108;  
Best Local Similarity 66.7%; Pred. No. 3.2e-20;  
Matches 54; Conservative 3; Mismatches 20; Indels 4; Gaps 3;  
QY 297 FNFC-LFEMESHVTOAGVQWPNLGSLOPLPGLKRFCSLSPSSWDYGHLPHPAN--F 353  
DB 14 FLFCFVEMESHVTOAGVQWMDLGSLOALPGFSPF-CLSLPSSWDYRCLPPSSANFFF 72  
QY 354 CIFIRGGVSPYLSGWSQFPDL 374  
DB 73 XIFXSXXVXPXPGWXQSPDL 93

RESULT 9  
US-09-800-729-193  
Sequence 193, Application US/09800729  
Patent No. 6605592  
GENERAL INFORMATION:  
APPLICANT: Ni et al.  
TITLE OF INVENTION: 32 Human secreted proteins  
FILE REFERENCE: PZ044P1  
CURRENT APPLICATION NUMBER: US/09/800,729  
CURRENT FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: PCT/US00/26013  
PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: 60/155,709  
PRIOR FILING DATE: 1999-09-24  
NUMBER OF SEQ ID NOS: 217  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 193  
LENGTH: 239  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-800-729-193

Query Match 12.5%; Score 254.5; DB 4; Length 239;  
Best Local Similarity 64.2%; Pred. No. 1.5e-19;  
Matches 61; Conservative 6; Mismatches 25; Indels 3; Gaps 3;  
QY 1 MEPSLLL-PRLECNCAISAHNRLPGSSDSPASASPVAGITGMCTHARLILYFVLVME 59  
DB 120 LWSLTLSPRLCSCSAISAHNRLPGSSNSPALASQVAGITGICHARQI-FVFLVETG 178  
QY 60 FLHVQAGLELPTSDPVSASQSAR-YRTGHAR 93  
DB 179 FCHVQAGLELLISGDSPASAFQSGAGIIGVSHR 213

RESULT 10  
US-09-800-729-190  
Sequence 190, Application US/09800729  
Patent No. 6605592  
GENERAL INFORMATION:  
APPLICANT: Ni et al.  
TITLE OF INVENTION: 32 Human secreted proteins  
FILE REFERENCE: PZ044P1  
CURRENT APPLICATION NUMBER: US/09/800,729  
CURRENT FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: PCT/US00/26013

PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: 60/155,709  
PRIOR FILING DATE: 1999-09-24  
NUMBER OF SEQ ID NOS: 217  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 190  
LENGTH: 310  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-800-729-190

Query Match 12.5%; Score 254.5; DB 4; Length 310;  
Best Local Similarity 64.2%; Pred. No. 2.2e-19;  
Matches 61; Conservative 6; Mismatches 25; Indels 3; Gaps 3;  
QY 1 MEPSLLL-PRLECNCAISAHNRLPGSSDSPASASPVAGITGMCTHARLILYFVLVME 59  
DB 191 LWSLTLSPRLCSCSAISAHNRLPGSSNSPALASQVAGITGICHARQI-FVFLVETG 249  
QY 60 FLHVQAGLELPTSDPVSASQSAR-YRTGHAR 93  
DB 250 FCHVQAGLELLISGDSPASAFQSGAGIIGVSHR 284

RESULT 11  
US-09-663-600A-114  
Sequence 114, Application US/09663600A  
Patent No. 6573068  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, Jean-Baptiste  
APPLICANT: Duclert, Aymeric  
APPLICANT: Bouqueleret, Lydie  
TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS  
FILE REFERENCE: 31 US3 CJP  
CURRENT APPLICATION NUMBER: US/09/663,600A  
CURRENT FILING DATE: 2000-09-15  
PRIOR APPLICATION NUMBER: 09/191,997  
PRIOR FILING DATE: 1998-11-13  
PRIOR APPLICATION NUMBER: 60/066,677  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/069,957  
PRIOR FILING DATE: 1997-12-17  
PRIOR APPLICATION NUMBER: 60/074,121  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/081,563  
PRIOR FILING DATE: 1998-04-13  
PRIOR APPLICATION NUMBER: 60/096,116  
PRIOR FILING DATE: 1998-08-10  
PRIOR APPLICATION NUMBER: 60/099,273  
PRIOR FILING DATE: 1998-09-04  
NUMBER OF SEQ ID NOS: 229  
SOFTWARE: Patent.pm  
SEQ ID NO 114  
LENGTH: 118  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIGNAL  
LOCATION: -83...-1  
NAME/KEY: UNSURE  
LOCATION: 28,32  
OTHER INFORMATION: Xaa = any one of the twenty amino acids  
US-09-663-600A-114

Query Match 12.4%; Score 253; DB 4; Length 118;  
Best Local Similarity 65.9%; Pred. No. 7.9e-20;  
Matches 60; Conservative 3; Mismatches 24; Indels 4; Gaps 3;  
QY 204 QSLNSVTOAGVQWPNLGSLOPLPGLKRFCSLSPSSWDYRPPRL-ANFFVFLVEMGPT 262  
DB 5 QSFTLVAQAGVQWRHLSSLQLLPPEFGFCSLSPSSWDYRPPPCPAGFFVFLVETGLH 64  
QY 263 MF--ARLILISGPCDLPASASQSAGITGVSH 291



```

; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: No. 6579710el Human Kinases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0281-USA
; CURRENT APPLICATION NUMBER: US/10/020,079
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255,103
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/289,422
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 876
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(876)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-020-079-30

```

```

Query Match      12.3%; Score 251; DB 4; Length 876;
Best Local Similarity 47.9%; Pred. No. 2.4e-18;
Matches 69; Conservative 9; Mismatches 38; Indels 28; Gaps 8;

Qy 211 QAGVQWRNIGSLQPLPPGPKLFSCPSLLSSWDYRR--PPRLANFFVFLVEMGFT-MEARL 267
Db 749 QAGVQWRDLGSLQPPPPRFKQFSCLSLPSWDYRHAPPHPAN-FVFLVETGFLHVEAGL 807

Qy 268 -ILISGPCDLPASASQAGITGVSHHARLIFNFCLEFEMESHVSQTQAGVQWPNLGSLOPLP 326
Db 808 ELLTSG--DLPASASQAGITGVSHRAQP--EVCEFN-RKHTGQR----- 847

Qy 327 PGLKRFSCLSLPSWDYGHLPHP 350
Db 848 ---EQWVCAGSERAWMRDLQGRP 868

```

Search completed: September 13, 2005, 19:38:54  
Job time : 44 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 15, 2005, 21:58:59 ; Search time 4030 Seconds  
(without alignments)  
3541.961 Million cell updates/sec

Title: US-09-380-203-2  
Perfect score: 2034  
Sequence: 1 MEFSLLPLRLCNGAISAH.....PIRGVSPYLSGWQTPLDLR 375

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 3423944 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_pool/US09380203/runat\_13092005\_170811\_11685/app\_query.fasta\_1.519  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LIST=45  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LOCAL  
-DOCALIGN=200 -THRESH=SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09380203 @CNG\_1\_1\_4352 @runat\_13092005\_170811\_11685 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_hic: \*  
4: gb\_est3: \*  
5: gb\_est4: \*  
6: gb\_est5: \*  
7: gb\_est6: \*  
8: gb\_gssi: \*  
9: gb\_gss2: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	935.5	46.0	622	CA437412	UI-H-DT0-
2	830.5	40.8	2587	AL157469	Homo sapi
3	801.5	39.4	549	BM996475	UI-H-DT0-
4	800	39.3	3016	BC068461	Homo sapi
5	722.5	35.5	1875	BC009270	Homo sapi
6	708	34.8	1842	BC011119	Homo sapi
7	698.5	34.3	4191	AL832992	Homo sapi
8	695	34.2	2786	BC035101	Homo sapi
9	671	33.0	4087	BC024593	Homo sapi

10	670	32.9	2097	3	CR614786	full-length
11	670	32.9	2777	3	HS0802772	AL390133 Homo sapi
12	662	32.5	3021	3	BC032537	Homo sapi
13	657	32.3	2874	3	BC047576	Homo sapi
14	654.5	32.2	990	2	AW600804	N26-20 Hu
15	646.5	31.8	3775	3	CR627381	Homo sapi
16	645	31.7	6056	3	CR627178	Homo sapi
17	642.5	31.6	2716	3	CR749867	Homo sapi
18	620.5	30.5	2076	3	BC035510	Homo sapi
19	611.5	30.1	3566	3	HS0805040	Homo sapi
20	610.5	30.0	3146	3	BC036603	Homo sapi
21	608.5	29.9	8213	3	HS0809169	Homo sapi
22	604.5	29.7	857	5	BQ437943	AGENCOURT
23	603	29.6	628	6	CD690152	EST6675 h
24	597.5	29.4	2044	3	BC037797	Homo sapi
25	596	29.3	2509	3	BC036220	Homo sapi
26	592.5	29.1	613	6	CD709038	EST25565
27	589.5	29.0	1345	1	AV762220	AV762220
28	587	28.9	1683	3	BC038849	Homo sapi
29	583.5	28.7	895	5	BU603620	AGENCOURT
30	582.5	28.6	5165	3	HS0802888	Homo sapi
31	581	28.6	811	5	BX643852	DKF2p7810
32	577.5	28.4	910	5	BQ722917	AGENCOURT
33	575	28.3	5797	3	CR749233	Homo sapi
34	573.5	28.2	3620	3	HS0804841	Homo sapi
35	573	28.2	946	5	BQ589903	AGENCOURT
36	571	28.1	725	7	CR789571	DKF2p4591
37	571	28.1	766	2	BF381650	601815637
38	570.5	28.0	935	5	BU838401	AGENCOURT
39	570	28.0	746	4	BG431326	602499916
40	570	28.0	1545	3	BC030737	Homo sapi
41	568	27.9	6146	8	AQ839831	260113-C4
42	567.5	27.9	737	1	AV700545	AV700545
43	567.5	27.9	767	1	AV700498	AV700498
44	566.5	27.9	3143	3	HS0805999	Homo sapi
45	565.5	27.8	1809	3	CR604283	full-length

ALIGNMENTS

RESULT 1  
CA437412  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

CA437412 622 bp mRNA linear EST 08-NOV-2002  
UI-H-DT0-avk-a-06-0-UI-s1 NCI CGAP DT0 Homo sapiens cDNA clone  
UI-H-DT0-avk-a-06-0-UI 3', mRNA sequence.

CA437412  
CA437412.1 GI:24801832

EST.  
Homo sapiens (human)

EST.  
Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (Bases 1 to 622)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgapsb@mail.nih.gov](mailto:cgapsb@mail.nih.gov)

Tissue Procurement: Dr. Jose Mercuende

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)

The following repetitive elements were found in this cDNA

sequence: 11-305, >ALU (matched complement) 221-440, >ALU (matched

complement) 452-620, >ALU (matched complement)

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers  
1..622  
/organism="Homo sapiens"

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-DT0-avk-a-06-0-UI"
/tissue_type="Metastatic Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP DT0"
note="Organ: Lung; Vector: pT773-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI_CGAP DT0 is a cDNA library containing the following
tissue(s): Metastatic Chondrosarcoma in Lung. The library
was constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT773-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(AGT)18 tail. The sequence tag for this library is
AACTGTTCCG.
TAG_TISSUE=Lung metastatic chondrosarcoma
TAG_LIB=UI-H-DT0
TAG_SEQ=AACTGTTCCG"

ORIGIN

Alignment Scores:
Pred. No.:          Length:      622
Score:             935.50      Matches:    189
Percent Similarity: 93.14%      Conservative: 1
Best Local Similarity: 92.65%    Mismatches: 6
Query Match:       45.99%       Indels:      2
DB:                6           Gaps:        2

US-09-380-203-2 (1-375) x CA437412 (1-622)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
DB 22 ATGGAGATTTTCGCTCTGTGCGCCAGGCTGGAGTGCAATGGCGCAATCTCAGCTCACCGC 81
QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerProValAlaGlyIle 40
DB 82 AACCTCGGCTCCCGGGTTACAGCGATTCTCTGCTCAGCTCCTCCTAGTAGTGGATT 141
QY 41 ThrGlyMetCysThrHisAlaArgLeuIleLeuTyrPhePheLeuValGluMetGluPhe 60
DB 142 ACAGGCATGTGCCACACCGCTCGGCTAAATTTGTATTTTTTTTAGTAGAGATGGAGTTT 201
QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAla 80
DB 202 CTCCTATGTTGGTCAGCTCGTCTCGAACTCCCGACCTCAGATGATCCTCCCGTCTCGGCC 261
QY 81 SerGlnSerAlaArgTyrArgThrGly-----HisHisAlaArg-LeuCysLeuAlaAs 98
DB 262 TCCCAAGTGCTGAGATT---ACAGGCATGAGCCACCACCTCCCGGCTCTGCTGGCTTAA 318
QY 98 nPheCysGlyArgAsnArgValSerLeuMet-CysProSerTrpSerProGluLeuLysG 118
DB 319 TTTTGTGTAGAAAACAGGGTTTCACTGATGTTGCCAACAGCTGCTCTCTGAGCTCAAGC 378
QY 118 InSerThrCysLeuSerLeuProLysCysTrpAspTyrArgArg--AlaAlaValProGl 137
DB 379 AGTCACCTGCTCAGCTCCTCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCTCGCTGG 438
QY 137 yLeuPheIleLeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGl 157
DB 439 CCTTTTATTTATTTTTTAAACACACAGGTTCCCACTCTTACCAGATGAATGCA 498
QY 157 nTrpCysAspHisSerSerLeuGlnProSerThrProGluIleLys-HisProProAlaAs 177
DB 499 GTGGGTGTATCATCAGCTCATCGAGCTTTCAACTCTCTGAGATCAAGCAATCTCTCGCT 558
QY 177 erAlaSerGlnValAlaGlyThrLysAspMetHisHisTyrThrTrpLeuIlePheIleP 197
```

```
Db 559 CAGCTCCCAAGTAGCTGGACCAAGACATGCACCACTACACCTGGCTAATTTTATTT 618
QY 197 he 197
Db 619 TT 620

RESULT 2
HSM802453/c
LOCUS HSM802453 2587 bp mRNA linear HTC 22-SEP-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp761N0823 (from clone DKFZp761N0823).
ACCESSION AL157469
VERSION AL157469.1 GI:7018485
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2587)
AUTHORS Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B., Amid,C.,
Osanger,A., Fobo,G., Han,M. and Wiemann,S.
SOURCE The German cDNA Consortium
DIRECT SUBMISSION
SUBMITTED (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
CLONE FROM S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp761N0823) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp761N0823
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES
Location/Qualifiers
1..2587
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp761N0823"
/tissue_type="amygdala"
/clone_lib="761 (synonym: hamy2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="RAB member RAS oncogene family-like 5, N-terminus
truncated"
gene 1..2587
/clone="DKFZp761N0823"
CDS <3..284
/clone="DKFZp761N0823"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH10693.1"
/db_xref="GI:50949389"
/translation="PSHREMEMWYSCFVQPSLQDTQCMLIAHHKPSGDDKGSLSL
SPPLNKLVLHNSLNLEDDPEETRMFIKLYKLSINSMSRSRDRENSIMT"

ORIGIN

Alignment Scores:
Pred. No.:          Length:      2587
Score:             830.50      Matches:    236
Percent Similarity: 49.34%      Conservative: 26
Best Local Similarity: 44.44%    Mismatches: 99
Query Match:       40.83%       Indels:     175
DB:                3           Gaps:        13

US-09-380-203-2 (1-375) x HSM802453 (1-2587)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20
DB 2566 ATGGAGATTTTG-CTCTGTGTTGCCAGGCTGGAGTGCAATGGCAATCTCGGCTTACTGC 2508
```

```
QY 21 AsnLeuArgLeuProGlySerSerPheSerProAlaSerAlaSerProValAlaGlyIle 40
Db 2507 AACCTCCGCTCCAGGTTACAGGATTCCTCTGCTCAGCTCCAGGTAGCTGGATT 2448
QY 41 ThrGlyMetCysThrHisAlaArgLeuLeuLeuTyrPhePheLeuValGluMetGluPhe 60
Db 2447 ACAGGCACGTGTACACGCCAGCTAATT---TTTGATATTTTAGTAGACGGGGTTT 2391
QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspSerProSerValSerAla 80
Db 2390 CACCATGTTGGACAGCTGCTTGAACCTCCTGACCTCAGGATTTCCACCCGCTCAGAC 2331
QY 81 SerGlnSerAla----- 84
Db 2330 TCCCAAGGTTGGGATTACAGCGTAAGCCACTGACCCCGCTCCAGGCTCAAGCAAT 2271
QY 85 ArgTyrArgThrGlyHisHisAlaArg-----LeuCysLeuAla 97
Db 2270 TCTTATGCCCTCAGC-CACATGAGTAGCTGTTATACAGGTGTGTGCCAACATGCTGGCG 2212
QY 98 AsnPheCys-----GlyArgAsnArgValSerLeuMetCysProSerTrp----- 112
Db 2211 AATTTTGTATTTTATGATAGATGGGGTTTCG-----CCATGTTGGCCAGGCTGT 2161
QY 113 -----SerProGluLeuLysGlnSerThrCysLeuSerLeuProLysCysTrpAspTyr 130
Db 2160 TCTTGAACCTCGGCTTCCATCAGTTCCTCCATCTAGCTCCCAAGTGTGGGATTAC 2101
QY 131 Arg-----ArgAlaAlaValProGlyLeuPheLeuPheLeuPheLeuArgHis 146
Db 2100 AGGTGTGACCCACTGCAACTGCCAGGCACTTCAATTTTTTTTTTTTTTTTGTCTGTCTCAC 2041
QY 147 ArgCysProThrLeuThrGlnAspGluValGlnTrpCysAspHisSerSerLeuGlnPro 166
Db 2040 -----CAGGCTGGAGTGCAGTGTGGTGGATCTCGGCTCACTGCAGTC- 2000
QY 167 SerThrProGluLeuLeuHisProProAlaSerAlaSerGlnValAlaGlyThrIysAsp 186
Db 1999 TCCGCTCGCGGTTTCAGGC-AATCTGCTCACTCCCTCTGAGTAGCTGGGACTACAGGC 1941
QY 187 MetHisHisTyrThrTrpLeuLeuPheLeuPheLeu----- 198
Db 1940 ATGCCCCACCATGCCAGCTAATTTTGTATTTTATGATTTTATGACAGACGGGGTTTCACCATGTT 1881
QY 198 ----- 198
Db 1880 GCCCAGGATGTCTCGATCTGACCTGTGATCGCTCGCTCGGCTCCCAAAGTGC 1821
QY 199 -----Phe-AsnPheLeu----- 202
Db 1820 TGGGATCACAGGCTGAGCCACCAGCCAGCCAGGCACTTCAATTTTTTGTCTCTGTG 1761
QY 202 ----- 202
Db 1760 TGTCTCGGAATAACTGAGAAAGTGCCACAGTATGATTTGGGGTTTACAAACATATTTTA 1701
QY 203 -----ArgGlnSerLeuAsn----- 207
Db 1700 GTGAGTAGGCAAAATCTCAAAATACAAATCTATGAATAAGGATCAAGTATACGTTTCAATT 1641
QY 207 ----- 207
Db 1640 GGCATTTTAAGTACCCAGGTTCCCTCATTTATCTGAGGAGGGCAGCATATGAACATTAT 1581
QY 208 -----SerValTh 210
Db 1580 TTAGACTTCTTCTAAACATTTAATGTAATTTAATTTATTTAGACAGGATGCTGTGTC 1521
QY 210 rGlnAlaGlyValGlnTrpArgAsnLeuGlySerLeuGlnProLeuProGlyPheIly 230
Db 1520 CCAGGCTGAGTATAGTGGCGGATCTCGGCTCAGTGAATCTCTGCTCCAGGCTCAA 1461
```

```
QY 230 sLeuPheSerCysProSerLeuLeuSerSerTrpAspTyrArgArgProPro-ArgLeuA 250
Db 1460 GCAATTTCTTATGCTCCAGCCAGGAGTAGCTGGTATCACAGGCTGGCCCAACACACC 1401
QY 250 laAsnPhePheValPheLeuValGluMetGly-PheThrMetPheAlaArgLeuLeuLeu 269
Db 1400 TGTAATTTTGTATTTTAGTAGACGGGATTTTGGCATGTTGGCCAGGCTGGTCTCG 1341
QY 270 lIleSerGlyPro--CysAspLeuProAlaSerAlaSerGlnSerAlaGlyIleThrGlyV 289
Db 1340 AACTCCCGCTCAAGTGTATCGGCTGCTCAGCATCCCAAGTCTGGGATTACAGTG 1281
QY 289 sIleSerHisHis-----AlaArgLeuIlePheAsnPheCysLeuPheGluMetGluSerH 307
Db 1280 TGAGCCACCGCTCTCGGCTAGAACTCTTTTTT-----TTTGAGATGGAGTCTT 1231
QY 307 sIleSerValThrGlnAlaGlyValGlnTrpProAsnLeuGlySerLeuGlnProLeuProP 327
Db 1230 GCTCTGTTGCCCAAGCTGGAGTACAGTGGCGCAATCTTGGCTCACTGCCAACCTCTGCCTC 1171
QY 327 roGlyLeuLysArgPheSerCysLeuSerLeuProSerSerTrpAspTyrGlyHisLeuP 347
Db 1170 CCGGCTCAACAATCTTATGCTCAGCTCCCAAGTAGCTGGATCACAGCAGCGCG 1111
QY 347 roProHisProAlaAsnPheCysIlePheIleArgGlyGlyValSerProTyrLeuSerG 367
Db 1110 CACCACACCGAGTATTTTCTATTTTAGTAGAGATGGGTTTTCACCACGTTGGTCAG 1051
QY 367 lyTrpSerGlnThrProAspLeu 374
Db 1050 GATGCTCGATCTCTGACCTC 1028

RESULT 3
LOCUS BM996475 549 bp mRNA linear EST 17-JUN-2002
DEFINITION UI-H-D70-avl-i-16-0-UI.s1 NCI_CGAP_D70 Homo sapiens cDNA clone
IMAGE:5881047 3', mRNA sequence.
ACCESSION BM996475
VERSION BM996475.1 GI:19721376
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 549)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
Cloning Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
The following repetitive elements were found in this cDNA
sequence: 11-305, >ALU (matched complement) 221-440, >ALU (matched
complement) 243-535, >SVA#Other
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Location/Qualifiers
1..549
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5881047"
/tissue_type="Metastatic Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_D70"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
```



modified polylinker: Site 1: EcoR I; Site 2: Not I;  
 NCI\_GCAP\_DT0 is a cDNA library containing the following  
 tissue(s): Metastatic Chondrosarcoma in Lung. The library  
 was constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into pT73-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (dT)18 tail. The sequence tag for this library is  
 AACTGTCGG.  
 TAG\_TISSUE=lung metastatic chondrosarcoma  
 TAG\_LIB=UI-H-DT0  
 TAG\_SEQ=AACTGTTCCG

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,45e-48 Length: 549  
 Score: 801.50 Matches: 165  
 Percent Similarity: 92.18% Conservative: 0  
 Best Local Similarity: 92.18% Mismatches: 6  
 Query Match: 39.41% Indels: 8  
 DB: 5 Gaps: 2

US-09-380-203-2 (1-375) x BM996475 (1-549)

QY 1 MetGluPheSerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArg 20  
 Db 22 ATGGAGTTTTCGCTCTGTGTCAGGCTGGAGTGCATGGCGCAATCTCAGCTCACC 81  
 QY 21 AsnLeuArgLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIle 40  
 Db 82 AACCTCCGCTCCCGGGTTCAAGCGATTCTCTGCTCAGCTCCCTAGTAGTGAGATT 141  
 QY 41 ThrGlyMetCysThrHisAlaArgLeuLeuLeuTyrPhePheLeuValGluMetGluPhe 60  
 Db 142 ACAGGATGTGCACACAGCTCGGCTAAATTTTGTATTTTTTTAGTAGATGGAGTTT 201  
 QY 61 LeuHisValGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAla 80  
 Db 202 CTCATGTTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCTCCGCTCTCGGCC 261  
 QY 81 SerGlnSerAlaArgTyrArgThrGly-----HisHisAlaArg-LeuCysLeuAlaAs 98  
 Db 262 TCCCAAAAGTGCTGAGATT---ACAGGCATGAGCCACCATGCCCGGCTCTGCTGGCTAA 318  
 QY 98 nPheCysGlyArgAsnArgValSerLeuMet-CysProSerTrpSerProGluLeuLysG 118  
 Db 319 TTTTGTGTAGANACAGGGTTTCACTGATGTGTGCCAGCTGGTCTCTGAGCTCAAGC 378  
 QY 118 InSerThrCysLeuSerLeuProLysCysTrpAspTyrArgArg--AlaAlaValProGl 137  
 Db 379 AGTCACCTGCTCAGCTCCCAAGTGTGGATTACAGGCGTGAGCCACCGCTGCTGG 438  
 QY 137 YleuPheHelleLeuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGl 157  
 Db 439 CCTTTTATTTATTTTTTTTAAAGACACAGGTGTCCCATCTTATCCAGGATGAAGTGCA 498  
 QY 157 nTrpCysAspHisSerSerLeuGlnProSerThr-ProGluIleLys 172  
 Db 499 GTGGTGTGATCAGCTCACTGAGCCITCACTCGCTGAGATCAAG 545

RESULT 4  
 BC068461/c  
 LOCUS BC068461 3016 bp mRNA linear HTC 19-JUL-2004  
 DEFINITION Homo sapiens cDNA clone IMAGE:30342539, with apparent retained  
 intron.  
 ACCESSION BC068461  
 VERSION BC068461.1 GI:46249759  
 KEYWORDS HTC.  
 SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 3016)

REFERENCE  
AUTHORS

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
 Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,  
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932

## TITLE

## JOURNAL

## PUBMED

2 (bases 1 to 3016)

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REMARK

## COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Dr. Stefan Hansson  
 cDNA Library Preparation: Michael Brownstein / Ted Usdin  
 Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 168 Row: m Column: 1  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 21361917  
 This clone has the following problem: retained intron.

## FEATURES

## source

1..3016  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30342539"  
 /tissue\_type="Placenta, normal"  
 /clone\_lib="NIH MGC\_147"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.53e-47 Length: 3016  
 Score: 800.00 Matches: 216  
 Percent Similarity: 57.51% Conservative: 29  
 Best Local Similarity: 50.70% Mismatches: 103  
 Query Match: 39.33% Indels: 81  
 DB: 3 Gaps: 11

US-09-380-203-2 (1-375) x BC068461 (1-3016)

```

Qy 4 SerLeuLeuProArgLeuGluCysAenGlyAlaIleSerAlaHisArgAenLeuArg 23
Db 2989 TCACCTCTTGTCCAGGCTAGAGTGCATAGTGGATCTTGGCTCAGTCAACATCCGC 2930

Qy 24 LeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIleThrGlyMet 43
Db 2929 CTCACGGGTTCAAGCAATCTCTCGCTCAGCTCAGGCTCAGAGTAGCTGGATTACAGGCATG 2870

Qy 44 CysThrHisAlaArgLeuLeuLeuTyPhePheLeuValGluMetGluPheLeuHisVal 63
Db 2869 CCCCACCCAGCTGGCTAATTTTGTG- - - - -TTTGTAGTAGAGTGGGGTTTGTGATGTT 2814

Qy 64 GlyGlnAlaGlyLeuGluLeuProThrSerAspSerProSerValSerAlaSerGlnSer 83
Db 2813 GGTCTGGCTGGTCTTGACTGCTGACCTCAGTGTCTGCCCGCTCGGCTCCCGAAGT 2754

Qy 84 Ala--ArgTyArgThrGlyHisAlaArgLeuCysLeuAlaAenPheCysGlyArgA 103
Db 2753 GCTGGGATTACAGG- - - - -CATGAGCCACCGCTGCTGGCTAGGTT- - - - - 2713

Qy 103 snArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThrCysLeuS 123
Db 2712 - - - - -CTTGACCTTTTGTAGGGCTAAACAGTCTTCAATAAACACAACTTACTTTTGTG 2658

Qy 123 erLeuProLysCysTrpAspTyArgArgAlaAlaValProGlyLeuPheLeuPheP 143
Db 2657 TT- - - - -TGTTTGTGTTGACAGAGT- - - - -TTT 2631

Qy 143 heLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAspHisSers 163
Db 2630 TTTTCAGACAG- - - - -ACAGCCAGCTGGAGGGAGTGGCATGATCTCAGCT 2583

Qy 163 erLeuGlnProSerThrProGluIleLys-HisProAlaSerAlaSerGlnValAla 182
Db 2582 CACTGTAACTCCACCTCCTGGGTTTCAAGTGAATCTTAGCGCTCAGCCTCCTGAGAAGCT 2523

Qy 183 GlyThrLysAspMetHisIstyrThrTrpLeuIlePheIle-Phe- - - - - 197
Db 2522 GGGATTACAGACATACACACACCTCGTGTAAATTTTGTATTTTGTAGTAGAGACAGGA 2463

Qy 198 - - - - -IlePheAenPheLeuArgGlnSerLeuAen- - - - - 207
Db 2462 TTTCCACCATGTTGGCCAGCTGCTTGAACCTCTGGCTCAGATGATCCACTCAGCTCG 2403

Qy 207 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - 207

Db 2402 GCCTCCCAAAATCTGGGATTACAGGCATGAGCCACGCGCGGCCCAACAACTTACT 2343

Qy 208 - - - - -SerValThrGlnAlaGlyValGlnTrpArgAs 218
Db 2342 TTTTGTGTGTGTGATGGAGTCTGCTCTGTACCCAGGCTGGAGTGCAGTGGCAAA 2283

Qy 218 nLeuGlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLe 238
Db 2282 TCTGGCTCACTGCAACCTCCGCTCTCTGGTTCAGCGATTCTTCTGCTCAGGCTCTCT 2223

Qy 238 uSerSerTrpAspTyArgArg- - -ProProArgLeuAlaAenPhePheValPheLeuVa 257
Db 2222 GAGTAGCTGGGAATACAGGACACACACACACGCTCAGCTAATTTTGTATTTTGTAGTAG 2163

Qy 257 1GluMetGlyPheThrMetPheAlaArgLeuIleLeuIleSerGlyProCysAspLeuPr 277
Db 2162 -AGACAGGATTCACCATGTTAGGAGGAGCGGTCTCAATCTCTGACCTTGTGATCCGCC 2104

Qy 277 oAlaSerAlaSerGlnSerAlaGlyIleThrGlyValSerHis- - - - -HisAlaArgLe 295
Db 2103 CGCCTTGGCTTCCAAAGTGTGGATGACAGGTGTGACCCACTGCCACTGCCACCACT 2044

Qy 295 uIlePhe- - - - -AsnPheCysLeuPheGluMetGluSerHisSerValThrGlnAl 312
Db 2043 TACTTTTAAAGTCATCTCTTTTTTTTTTTTTTTTTTTTGTGCTGTCTGTGCCCGGCG 1984

```

```

Qy 312 aGlyValGlnTrpProAenLeuGlySerLeuGlnProLeuProGlyLeuLysArgPh 332
Db 1983 TGGAGTCAGTGGGCAATCTTGGCTCACTGCAACCTCAGCTTCCAGGTTCCAGCTATT 1924

Qy 332 eSerCysLeuSerLeuProSerSerTrpAspTyArgGlyHisLeuProProHisProAlaAs 352
Db 1923 CTCCTGCC- -TCCATGTCAGTAGCTGGGATTACGGGTGCGGCCACCATGCTGCTGCTAA 1866

Qy 352 nPheCysIlePheIleArgGlyValSerProTyTrpLeuSerGlyTrpSerGlnThrPr 372
Db 1865 TTTTGTATTTTGTAGTAGAGATGGGGTTTCCACCATATTGGCCAGGCTGCTCTCAAACTCC 1806

Qy 372 oAspLeuArg 375
Db 1805 TGACCTCAAG 1796

RESULT 5
BC009270
LOCUS
DEFINITION BC009270 1875 bp mRNA linear HTC 04-MAR-2003
IMAGE:3342760, mRNA.
ACCESSION BC009270
VERSION BC009270.1 GI:14627272
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1875)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stort,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LiNL at: http://image.llnl.gov
Series: JRAL Plate: 22 Row: j Column: 8
This clone has the following problem: retained intron.
FEATURES
location/Qualifiers
source
1..1875
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3342760"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH MGC 16"
/lab_host="DH10B-R"
/note="vector: pOTB7"
ORIGIN
Alignment Scores: 3.69e-42 Length: 1875
Pred. No.: 722.50 Matches: 193
Score:

```

Percent Similarity:	52.82%	Conservative:	13
Best Local Similarity:	49.49%	Mismatches:	62
Query Match:	35.52%	Indels:	127
DB:	3	Gaps:	7

  

US-09-380-203-2 (1-375) x BC009270 (1-1875)			
Qy	4	SerLeuLeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArgAsnLeuArg	23
Db	576	TCGCTTTTGTTCGCCAGGCTGGAGTGCAATGGCTCGATTTTCAGCTCACTGCAACCTCTGC	635
Qy	24	LeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIleThrGlyMet	43
Db	636	CTGCCGGGTTCAAGCGATTCTCTGCCTCAGCCTCCGAGTAGCTGGGATTACAGGCATG	695
Qy	44	CysThrHisAlaArgLeuLeuLeuTyrrPhePheLeuValGluMetGluPheLeuHisVal	63
Db	696	TGCCACCATGCCCGGCTAAATTTTGTG---CTTTTAGTAGAGACGAGGTTTCTCCATGTT	751
Qy	64	GlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAlaSerGlnSer	83
Db	752	GGTCAGGCTGCTTTGAACCTCCGACCTCAGGTGATCCACC-----	792
Qy	84	AlaArgTyrrArgThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCysGlyArgAsn	103
Db	792	-----	792
Qy	104	ArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThrCysLeuSer	123
Db	793	-----TGCTTCAGC	801
Qy	124	LeuProLysCysTrpAspTyrrArgAlaAlaValProGlyLeuPheIleLeuPhePhe	143
Db	802	CTCCCAAGTCTGGGATTACAGG-----	825
Qy	144	LeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAspHisSerSer	163
Db	826	-----CATGA-GCC	833
Qy	164	LeuGlnProSerThrProGluIleIleYHisProProAlaSerAlaSerGlnValAlaGly	183
Db	834	ACCATGCTGGCAGCT-----	848
Qy	184	ThrLysAspMetHisHisTyrrThrTrpLeuIlePheIlePheAsnPheLeuArg	203
Db	849	-----GGTTTTTGTGAGA	860
Qy	204	GlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGlySerLeuGln	223
Db	861	TGGAATCTCGT-TATGTCAACCCAGCTGGAGGCGAGTGGCAGCATCTCAGCTCACTGTAA	919
Qy	224	ProLeuProProGlyPheLysLeuPheSerCysProSerLeuLeuSerSerTrpAspTyr	243
Db	920	CTTTTGGCGGT---GGGTTTCAAGCGATTCTCTGCGCTCAGCCTCCCAAGTAGCTGGGATTAC	977
Qy	244	ArgArg---ProProArgLeuAlaAsnPhePheValPheLeuValGluMet-GlyPheTh	262
Db	978	AGGCAGCTCCACACACACCCCGCTAAATTTTTTTGTATTTTTTAGTAGAGATGGGGTTTCAC	1037
Qy	262	rMetPheAlaArgLeuIleLeuIleSerGlyPro--CysAspLeuProAlaSerAlaSer	281
Db	1038	CATATTGGCCAGGCTGGTCTTGAACCTCTGACCTCAAGTGATCCGCTGCCTTCGCTCC	1097
Qy	282	GlnSerAlaGlyIleThrGlyValSerHisHisAlaArg-----LeuIlePhe	297
Db	1098	CAAAAGTCTGGGATTACAGGCGTGAGCCACCGTGCCTCCCATCAACTTAACCTTTTAATTTT	1157
Qy	298	Asn-----PheCysLeuPheGluMetGluSerHis	307
Db	1158	TCTCATGTAAATTTGTTCAATTTCTTTTTTTTTTTTTTTTTTTTTTTTGTGACAACTC	1217
Qy	308	SerValThrGlnAlaGlyValGlnTrpProAsnLeuGlySerLeuGlnProLeuProPro	327

```

QY 6 LeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArgAsnLeuArgLeuPro 25
Db 598 CTGTCGCCAGGCTGAGTGCATCTCGCTTACTGCAAGCTCCGCTCCCG 657
QY 26 GlySerSerAspSerProAlaSerProValAlaGlyTleThrGlyMetCysThr 45
Db 658 GGTTCACACTTCTCTCGCTCAGCTCCAGTAGCTGGAGTACTGGTCCGCCAC 717
QY 46 HisAlaArgLeuLeuLeuPhePheLeuValGluMetGluPheLeuHisValGlyGln 65
Db 718 CAGCGCCGCTAATTTT-TTCTATTTTATAGACGGGTTTCCACGCTGTAGCCAG 776
QY 66 AlaGlyLeuGluLeuProThr-SerAspSerProSerValSerAlaSerGlnSerAlaAr 85
Db 777 GATGGTCTTGTGATCTGCTGACCTGTGATTCACCTGCC---TCGGCTCCCAAGTGTGG 833
QY 85 gTyrArgThrGly-----HisAlaArg-LeuCysLeuAlaAsnPheCysGlyArgA 103
Db 834 GATT---ACAGGAGTGAGCCAGCGCCGGCTTATTTTA-----TTTTTAATAGAG 884
QY 103 snArgValSerLeuMetCysPro-----SerTrpSerProGluLeuLysGlnSerThrC 121
Db 885 AT---GTCTCATATGTCGGGACTAGTCTCAAACTCTGGGCTCAAGCATCTCTCC 941
QY 121 ysLeuSerLeuProLysCysTrpAspTyrArgArg--AlaAlaValProGlyLeuPheI 140
Db 942 A-CTTGGCTCCCAAGTCTGGGATTACAGTTCTGAGCAGTGTGCTGGCTCTCTTT-- 998
QY 140 eleuPhePheLeuArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAs 160
Db 999 ----TTTTTTTGTAGATA-GAATCTCCCTGTGTACCCAGGTGTGACGTGTGGCA 1053
QY 160 phiSerSerLeu-----GlnProSerThrProGluIleLys-H 173
Db 1054 TCTAGCTCACTGCACCTCGCTAGTTCAGCGGCACTCGTCTCTCGGTTCAAGCA 1113
QY 173 isProProAlaSerAlaSerGlnValAlaGlyThrLysAspMethHisIstYrThrTrpL 193
Db 1114 ATTCTCTCGCTCAGCTCCGAGTAGCTGAGATTACAGGCGCATGTCACCCAGCTGCG 1173
QY 193 euilePheIlePheIlePheAsnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaG 213
Db 1174 TAAATTTTGTGTTTTTTTGTGTTGTGAGACAGAGTCTCTC---TCTGTAC-CAGGCTG 1229
QY 213 lyValGlnTrpArgAsnLeuGlySerLeuGlnProLeuProProGlyPheLysLeuPheS 233
Db 1230 GAGTGCAGTGGGTGATCTGGCTCACTGCAACCTCCGCTCCCGGTTTACGCCATTCT 1289
QY 233 erCysProSerLeuLeuSerSerTrpAspTyrArgArgProPro-ArgLeuAlaAsnPhe 252
Db 1290 CTGTGCCAGC-CTCTCAAGTAGCTGGACTACAGGCGCGCCGCCATCGCCGGCTAATT 1348
QY 253 PheValPheLeuValGluMet-GlyPheThrMetPheAlaArgLeuLeuLeuIleSerG 272
Db 1349 TTTGTATTTTATAGTAGTGGGTTTCACTATGTTGGCCAGGATGGTCTCAAACTCTG 1408
QY 272 yProCysAspLeuProAlaSerAlaSerGlnSerAlaGlyIleThrGlyValSerHisH 292
Db 1409 ACCTCATGATTCACCACTCAGCTCCCAAACTGCTGGATTACAGGTTGAGGCCATCA 1468
QY 292 sAlaArgLeuIlePheAsnPheCysLeuPheGluMetGluSerHisSerValThrGlnAl 312
Db 1469 CGCCGAGCTTTTIT-----TTTTTTTTTTAGATCTACCTCTGTGTGCCAGG 1519
QY 312 aGlyValGlnTrpProAsnLeuGlySerLeuGlnProLeuProProGlyLeuLysArgPh 332
Db 1520 GGGAGTCCAGTGGCGGTGATCTTTGGTCTCACTGCAACCGCGCTCCCGGTTTCAAGCAATT 1579
QY 332 eSerCysValSerLeuProSer 339
Db 1580 CTCCTGCCTCAGTCTCCCGAGT 1601

```

RESULT 7

```

HSM804303/c
LOCUS Homo sapiens mRNA; cDNA DKFZp666K083 (from clone DKFZp666K083).
ACCESSION AL832992
VERSION AL832992.1 GI:21733581
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 4191)
AUTHORS Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,
Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
CONSTRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp666K083) is available at the RZPD Deutsches
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp666K083
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.
FEATURES
location/Qualifiers
1..4191
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp666K083"
/db_xref="taxon:9606"
/clone="DKFZp666K083"
/tissue_type="stomach"
/clone_lib="666 (synonym: hato2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="genomic"
1..4191
/gene="DKFZp666K083"
2517..2675
/gene="DKFZp666K083"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH56307.1"
/db_xref="GI:52545758"
/translation="MMTRQVSKTSTHRSFLFLEVTGGCAASKHKRPGMVAHACNPFSTL
GGQGQIR"
ORIGIN
Alignment Scores:
Pred. No.: 5,55e-40 Length: 4191
Score: 698.50 Matches: 192
Percent Similarity: 50.11% Conservative: 32
Best Local Similarity: 42.95% Mismatches: 92
Query Match: 34.34% Indels: 135
DB: 3 Gaps: 11
US-09-380-203-2 (1-375) x HSM804303 (1-4191)
QY 6 LeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArgAsnLeuArgLeuPro 25
Db 2876 CTGTCCGCTGAGTGGAGTGCAGGCGTGCATCTCGCTCACTGCAACCTCCGCTCCCG 2817
QY 26 GlySerSerAspSerProAlaSerAlaSerProValAlaGlyIleThrGlyMetCysThr 45
Db 2816 GGTTCACACTTCTCTCGCTCAGCTCCCAAGTAGTGGACTACAGGCGCGTGCAC 2757
QY 46 HisAlaArgLeuLeuLeuPhePheLeuValGluMetGluPheLeuHisValGlyGln 65
Db 46 HisAlaArgLeuLeuLeuPhePheLeuValGluMetGluPheLeuHisValGlyGln 65

```

2756 TACACCCGGCTAAATTTCTGGCAATTTTAAATAGACAGCGGTTTCGCCATGCTTGGCCAG 2697  
QY  
66 AlaGlyLeuGluLeuProThrSerAspAspProSerValSerAlaSerGlnSerAlaArg 85  
Db  
2696 GCTGATCTCAAACTCCTGGCCTCATCTGAT----- 2667  
QY 86 TyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCysGlyArgAsnArgVal 105  
Db  
2667 ----- 2667  
QY 106 SerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThrCysLeuSerLeuPro 125  
Db  
2666 -----CTGCC-CCC-----CTTGGCCTCCCA 2647  
QY 126 LysCysTrpAspTyrArgArg--AlaAlaValProGlyLeuPheLeuPheLeuPheLeu 145  
Db  
2646 AAGTGTGGGATACAGGCGTGAGCCACCATCCCGCGCTCCCTTATGTTTGTATGCGAG 2587  
QY 145 rgHis----- 149  
Db  
2586 CACATCCTCCAGTAATCTTAAGACAGCGCTTCTGTGAGTTGAGGTTTGTAGACCTGTTC 2527  
QY 149 roThrLeuThrGlnAspGluValGlnTrpCysAspHisSerSerLeuGlnProSerThrP 169  
Db  
2526 TAGTCATCATGACAGATCTTTCATGGGGGAG-----TCATGTC 2485  
QY 169 roGluLeuLysHisProProAlaSerAlaSerGlnVal----- 181  
Db  
2484 CATCAGTCATTAATTAATCAAAATGTTCTCTCTTACTGTCTTAATAACTTCTTCTACTCTT 2425  
QY 182 -----AlaGlyThrLysAspMetHisHisTyrThrTrp----- 192  
Db  
2424 CCTTCATCTTTTCTTGGGAGCAGCATG-TTGAACCTCTCTGAATGGTCACTAATT 2366  
QY 193 -----LeuLeuPheLe----- 196  
Db  
2365 TCATTTCTCTTCTCATCTTTTGTGTTTAACTCTTGTGTTTGTGAATATCTTCTG 2306  
QY 197 -----PheLeuPheAsnLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyV 214  
Db  
2305 GGACCTTTTCTTACCTTTTGTGAGCGGAGTCTCGTCTATTGCCAGG-----C 2252  
QY 214 aGlnTrpArgAsnLeuGlySerLeuGlnProLeuProProGlyPheLysLeuPheSerC 234  
Db  
2251 TGAAGTGGCACAATCTGGGATCACTGCAAGCTCCACCTCTCGGTTCAGCCATTTCTCT 2192  
QY 234 ysProSerLeuLeuSerSerTrpAspTyrArg---ArgProProArgLeuAlaAsnPheP 253  
Db  
2191 ATGTCTAGCTCCGGAGCAGCTGGGACACAGAGTGGCCGCCACCGCGCCC-GCCCAATTTT 2133  
QY 253 heValPheLeuValGlu-MetGlyPheThrMetPheAlaArgLeuLeuLeuLeuSerGly 272  
Db  
2132 TTGTATTTTGTAGAGACAGGTTTCACTGTGTAGCCAGGATGCTCTCAATCTCCTGA 2073  
QY 273 ProCysAspLeuProAlaSerAlaSerGlnSerAlaGlyLeuThrGlyValSerHisHis 292  
Db  
2072 CCTCGTGATTG-----GCCCTCCCAAGTGTGGGATTACAGGCAATGAGCCACAGC 2023  
QY 293 Ala-ArgLeuLeuPheAsnPheCysLeu----- 301  
Db  
2022 ACCTGGCCTTTTCTTACTTTTATCTTTCAACACTCTATTGAATTTTAAAGATATA 1963  
QY 302 -----PheGluMetGluSerHisSerValThrG 311  
Db  
1962 TAAATAACATAGTTGTTTGTGTTTCTTTTGTACATACAGTCTCACTCTTTCACCCA 1903  
QY 311 nAlaGlyValGlnTrpProAsnLeuGlySerLeuGlnProLeuProProGlyLeuLysAr 331  
Db  
1902 AGCTGGAGTGCAGTGGTGTGATCTTGGCTCACTGCAACCTCACTCCCGGGTTCAAGCA 1843  
QY 331 gPheSerCysLeuSerLeuProSerSerTrpAspTyrGlyHisLeuProProHisProAl 351  
Db  
1842 ATTCTCCTGCTCAGCCTCCCAAGTAGCTGGGATTACAGGCGCTTGCACACCGCCGGC 1783

QY 351 aAsnPheCysLeuPheLeuArgGlyValSerProTyrLeuSerGlyTrpSerGlnTh 371  
Db 1782 TAAATTTTGTATTTTGTAGTAGAGACGGGTTTCTCCATGTTGGCCAGGCTGCTCTCGAAC 1723  
QY 371 rProAspLeuArg 375  
Db 1722 TCCTGACCTCAAG 1710  
RESULT 8  
LOCUS BC035101 2786 bp mRNA linear HTC 04-MAR-2003  
DEFINITION Homo sapiens, Similar to neuronal thread protein, clone  
IMAGE:5262055, mRNA.  
ACCESSION BC035101  
VERSION BC035101.1 GI:23272462  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2786)  
TITLES Direct Submission  
AUTHORS Strausberg,R.  
JOURNAL Submitted (31-JUL-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
contact: [amadane@systemsbiology.org](mailto:amadane@systemsbiology.org)  
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketterman, Anuradha  
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 73 Row: F Column: 21  
This clone has the following problem: retained intron.  
FEATURES  
Location/Qualifiers  
source  
1..2786  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5262055"  
/tissue\_type="Brain, hippocampus"  
/clone\_lib="NIH MGC\_95"  
/lab\_host="DH10B"  
/note="Vector: pBluescript"  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.02e-40 Length: 2786  
Score: 695.00 Matches: 184  
Percent Similarity: 50.24% Conservative: 23  
Best Local Similarity: 44.66% Mismatches: 77  
Query Match: 34.17% Indels: 131  
DB: 3 Gaps: 9  
US-09-380-203-2 (1-375) x BC035101 (1-2786)  
QY 8 ProArgLeuGluCysAsnGlyAlaIleSerAlaHisArgAsnLeuArgLeuProGlySer 27  
Db 1371 CCCAGGCTGGAGTGCAGTGCGCACATCTTGGCTCATTTGCTCACTTCCACCTCCAGGTTCA 1430  
QY 28 SerAspSerProAlaSerAlaSerProValAlaGlyIleThrGlyMetCysThrHisAla 47  
Db

Db 1431 AGTGATTCCTCGCTCAGCTCCCAAGTAGCTGGGATTACAGGCGTACGCCACCATGCC 1490  
QY 48 ArgLeuIleLeuTyrPhePheLeuValGluMetGluPheLeuHisValGlyGlnAlaGly 67  
Db 1491 CAGTTA---ACITTTGTATTTTAGTAGACGGCGTTTCATCATGTTGGCTAGGCTGT 1547  
QY 68 LeuGluLeuProThrSerAspAspProSerValSerAlaSerGlnSerAlaArgTyrArg 87  
Db 1548 CTTGAAC----- 1555  
QY 88 ThrGlyHisHisAlaArgLeuCysLeuAlaAsnPheCysGlyArgAsnArgValSerLeu 107  
Db 1555 ----- 1555  
QY 108 MetCysProSerTrpSerProGluLeuLysGlnSerThrCysLeuSerLeuProLysCys 127  
Db 1556 -----CTGACCTCAAGTATCGCCACCTTGGTCTTCCAAAGTAC 1597  
QY 128 TrpAspTyrArg---ArgAlaAlaValProGlyLeuPhePheLeuPheLeuArgHis 146  
Db 1598 TGGGACTACAGGCATGAGAAACCAAGCCGGCCATCTTTCT-CTTTTTFACCTG----- 1650  
QY 147 ArgCysProThrLeuThrGlnAspGluValGlnTrpCysAspHisSerSerLeuGlnPro 166  
Db 1651 -----AATTCAAGCTTTTATTCTTCTTCCATCTCTCTCTCTCTC 1692  
QY 167 SerThrProGluIleLysHisProProAlaSerAlaSerGlnValAlaGlyThrLysAsp 186  
Db 1693 TCT----- 1695  
QY 187 MetHisHisTyrThrTrpLeuIlePhePheLeuPheAsnPheLeuArgGlnSerLeu 206  
Db 1696 -----TTTTT-TTTTGAGACAGTCT--- 1715  
QY 207 AsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGlySerLeuGlnProLeuPro 226  
Db 1716 CACTATGTCACCCAGCTCGAGTGCAGTGGCGCAATCTCGGCTCACTGCAGCCTCTACCT 1775  
QY 227 ProGlyPheLeuLeuPheSerCysProSerLeuLeuSerSerTrpAspTyrArg---Arg 245  
Db 1776 CCAGATTCAAGGATTCCTCGCTCAGCCTCGAGTAGTGGGACTACAGGTGCGGTG 1835  
QY 246 ProProArgLeuAlaAsnPhePheValPheLeuValGluMet-GlyPheThrMetPheAl 265  
Db 1836 CCACCATGCGCGCTAAATTTTGTATTTTAGTAGATGGGTTCACCATTTTGGC 1895  
QY 265 aArgLeuIleLeuLeuSerGlyProCysAspLeuProAlaSerAlaSerGlnSerAlaG 285  
Db 1896 CAGGATGGTCTCAATCTCTGACCTCGTGTATCTGCTGGCTGCCAAAGTGTGG 1955  
QY 285 YleThrGlyValSerHisHisAla-ArgLeuIlePheAsnPheCysLeuPheGluMet- 304  
Db 1956 GATCAGGACGATGAACACCGCTCGGCTCTTTCTTTCTTTCTGTTCTTTCCATCT 2015  
QY 304 ----- 304  
Db 2016 GCGTCCCAACATCTCCCTACCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCT 2075  
QY 305 -----GluSerHis 308  
Db 2076 TTCT 2135  
QY 308 exValThrGlnAlaGlyValGlnTrpProAsnLeuGlySerLeuGlnProLeuProG 328  
Db 2136 CTGTTGCCAGGATGAATGCAAGTGCAGTGGCAAACTTAGTCTCACTACAACTCTGCTCCCA 2195  
QY 328 lLeuLysArgPheSerCysLeuSerLeuProSerSerTrpAspTyrGlyHisLeuPro 348  
Db 2196 GGTTCAGCAATCT 2255  
QY 348 roHisProAlaAsn-----PheCysIlePheIle-ArgGlyGlyValSerProTyrLeu 365  
Db 2256 CACACCAAGTAAGTTTTTTTTTTTATTTTATTTTATTTATTTATTTATTTATTTATTT 2315

QY 366 SerGlyTrpSerGlnThrProAspLeuArg 375  
Db 2316 CCAAACTGTCGTGAACCTCCTGGGCTCAAG 2345  
RESULT 9  
LOCUS BC024593  
DEFINITION Homo sapiens cDNA clone IMAGE:3914314, with apparent retained intron.  
ACCESSION BC024593  
VERSION BC024593.1 GI:22137609  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 4087)  
AUTHORS Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzyzanski,M.I., Skalska,U., Smalish,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
PUBMED 12477932  
REFERENCE 2 (bases 1 to 4087)  
AUTHORS Strausberg,R.  
TITLE Direct Submission  
JOURNAL Submitted (01-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www.shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcdpaxil.stanford.edu](mailto:mcdpaxil.stanford.edu)  
R. M.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 22 Row: m Column: 14  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis  
This clone has the following problem: retained intron.  
FEATURES  
source Location/Qualifiers  
1. 4087  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3914314"





/db\_xref="taxon:9606"  
/clone="CS0DF006YD23"  
/tissue\_type="Fetal brain"  
/plasmid="PCWVSPORT\_6"

## ORIGIN

## Alignment Scores:

Pred. No.: 2,79e-38 Length: 2097  
Score: 670.00 Matches: 177  
Percent Similarity: 52.53% Conservative: 20  
Best Local Similarity: 47.20% Mismatches: 77  
Query Match: 32.94% Indels: 103  
DB: 3 Gaps: 5

US-09-380-203-2 (1-375) x CR614786 (1-2097)

QY 6 LeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArgAsnLeuArgLeuPro 25  
DB 426 CTGTGCGCCCAAGTCGAGTGCAGTGGCTTGGCTCACTGCAACCTCCACCTCTCTG 485  
QY 26 GlySerSerAspSerProAlaSerAlaSerProValAlaGlyIleThrGlyMetCysThr 45  
DB 486 GTTTCAGGTATCTCTGCTCAGCTTCTGAGTGTGGATTATAGCCGATGCCAC 545  
QY 46 HisAlaArgLeuLeuLeuThrPhePheLeuValGluMetGluPheLeuHisValGlyGln 65  
DB 546 CATGCTGGCTAATT---TTTGTATTTTAAATAGAGAGCGAGTTCACCATCTGGTCAG 602  
QY 66 AlaGlyLeuGluLeuProThrSerAspProSerValSerAlaSerGlnSerAlaArg 85  
DB 603 GCTGTGTTTGAATCTCTGACCTCAATGATCCACCTGCTCTCTCCGAAAGCGTGG 662  
QY 86 TyrArg---ThrGlyHisAlaArgLeuCysLeuAlaAsnPheCysGlyArgAsnArg 104  
DB 663 ATTAGAGGTATGACGACCATGCCGAGGCTCTCTCT----- 701  
QY 105 ValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThrCysLeuSerLeu 124  
DB 701 ----- 701  
QY 125 ProLysCysTrpAspTrpArgArgAlaAlaValProGlyLeuPheIleLeuPheLeu 144  
DB 701 ----- 701  
QY 145 ArgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAspHisSerSerLeu 164  
DB 701 ----- 701  
QY 165 GlnProSerThrProGluIleLysHisProProAlaSerAlaSerGlnValAlaGlyThr 184  
DB 701 ----- 701  
QY 185 LysAspMetHisIleThrTrpLeuIlePhePheIlePheAsnPheLeuArgGln 204  
DB 702 -----TTTAAACCTCTTAAATATTTCTTTT---TTTTTGAGACGG 743  
QY 205 SerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGlySerLeuGlnPro 224  
DB 744 AGTCGCG---TCGTGTTCAGGCTGAGTGCAGTGGTGCATCTTGAAGCACTGTACT 802  
QY 225 LeuProProGlyPheLysLeuPheSerCysProSerLeuLeuSerSerTrpAspTrpArg 244  
DB 803 CTCACCTCTGGATTCAAGGATTCCTGCTCAGCTCTGAGTGTGGATTACAGG 862  
QY 245 ArgProPro-ArgLeuAlaAsnPhePheValPheLeuValGln-MetGlyPheThrMetP 264  
DB 863 CACCCACCAACGAGCCCACTAATTTTGTATTGTAGTACAGACAGGGTTTCACCATGT 922  
QY 264 healaArgLeuLeuLeuLeuSerGlyPro--CysAspLeuProAlaSerAlaSerGlnSe 283  
DB 923 TGGACAGGCTGTCTTGAAGCGCTGACCTCAGGTATTCCACCCCGCGCTCCCAAG 982  
QY 283 rAlaGlyIleThrGlyValSerHisAlaAlaArgLeuIlePheAsnPheCysLeuPheGln 303

DB 983 TGCTGGGATTACAGCGGTGAGCCACTGGCGCTG---GCTTATTTTATTC---TTTTTGA 1038  
QY 303 uMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGlySerLeuGln 323  
DB 1039 GATGGAGTCTTGTCTGTGTGCCAGGCTGGAATGAGTGGCGTGTCTTGGCTCACTGCA 1098  
QY 323 nProLeuProProGlyLeuLysArgPhe-SerCysLeuSerLeuProSerSerTrpAspTr 343  
DB 1099 ACCTCTGCTCTCCAGGTTCAAGTGATTCTTCAGCCTCAGGCTCCAGTAGCTGGATT 1158  
QY 343 yrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGlyGlyValSerP 363  
DB 1159 ACAGGTGCCACCAACCAACCACTAATTTTGTATTTTAGTAGAGACGGGGTTTCAT 1218  
QY 363 roTyrLeuSerGlyTrpSerGlnThrProAspLeu 374  
DB 1219 CATGTTGCCAGGCTGGTCTTGAACCTCCTGACCTC 1253  
RESULT 11  
LOCUS HSM802772/c 2777 bp mRNA linear HTC 22-SEP-2004  
DEFINITION Homo sapiens mRNA; cDNA DKFZp5471147 (from clone DKFZp5471147).  
ACCESSION AL390133  
VERSION AL390133.1 GI:9368829  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2777)  
AUTHORS Anorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,  
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M., and  
Wiemann,S.  
CONSRMT The German cDNA Consortium  
TITLE Direct Submission  
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764  
COMMENT Neuherberg, GERMANY  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
This clone (DKFZp5471147) is available at the RZPD Deutsches  
Reisourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.  
Please contact RZPD for ordering:  
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp5471147  
Further information about the clone and the sequencing project is  
available at http://mips.gsf.de/projects/cdna/.  
FEATURES  
Location/Qualifiers  
1..2777  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp5471147"  
/tissue\_type="brain"  
/clone\_lib="547 (synonym: hfbr1). Vector pSport1; host  
DH10B; sites NotI + SalI"  
/dev\_stage="fetal"  
/note="genomic"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.94e-38 Length: 2777  
Score: 670.00 Matches: 180  
Percent Similarity: 51.17% Conservative: 16  
Best Local Similarity: 47.00% Mismatches: 77  
Query Match: 32.94% Indels: 111  
DB: 3 Gaps: 11  
US-09-380-203-2 (1-375) x HSM802772 (1-2777)  
QY 1 MetGluPheSerLeuLeuLeuPro---ArgLeuGluCysAsnGlyAlaIleSerAlaHis 19





Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www.shgc.stanford.edu>  
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 90 Row: d Column: 9  
 This clone has the following problem: frame shifted.

#### FEATURES

Location/Qualifiers  
 1..2874  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4814349"  
 /tissue type="Brain, hippocampus"  
 /clone\_lib="NIH MGC\_95"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript"

#### ORIGIN

Alignment Scores:  
 Pred. No.: 3,63e-37 Length: 2874  
 Score: 657.00 Matches: 191  
 Percent Similarity: 54.16% Conservatives: 24  
 Best Local Similarity: 48.11% Mismatches: 103  
 Query Match: 32.30% Indels: 82  
 DB: 3 Gaps: 10

US-09-380-203-2 (1-375) x BC047576 (1-2874)

QY 4 SerLeuLeuLeuProArgLeuGluCysAenGlyAlaAlaSerAlaHis-ArgAsnLeuAr 23  
 Db 67 TGGCTCTCATGCGCGGCTGTAGTGCAGTGTGGCATCTCGCTCACCTTGAACCTCG 126  
 QY 23 gLeuProGlySerSerAspSerProAlaSerAlaSerProValAlaGlyIleThrGlyWe 43  
 Db 127 CTTCTGGGTTCAAGCGATCTCCTGCCTCGGCTCCCGAGTAGCTGGGATTACAGCGT 186  
 QY 43 tCysThrHisAlaArgLeuLeuLeuTyrrPhePheLeuValGluMetGluPheLeuHisVa 63  
 Db 187 GTGCCACCACACCCAGTAAATTTGTATTTTGTATAGTAGAGACGGGGTTTACCGTGT 246  
 QY 63 lGlyGlnAlaGlyLeuGluLeuProThrSerAspAspProSerValSerAlaSerGlnSe 83  
 Db 247 TGGTCAGGCTGCTTGAAC----- 267  
 QY 83 xAlaArgTyrrGthrGlyHisAlaArgLeuCysLeuAlaAsnPheCysGlyArgAs 103  
 Db 267 ----- 267  
 QY 103 nArgValSerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThrCysLeuSe 123  
 Db 268 -----CCTGATCTCAGGTGATCCACCGCTTGG 296  
 QY 123 rLeuProLysCysTrpAspTyrrArgArg--AlaAlaValProGlyLeuPheIleLeuPhe 142  
 Db 297 TCTCCAGAGTGTGGGATTGACGAGTAGTGCACCTGCGCTGGCCAAATTTGTGTTTT 356  
 QY 143 PheLeuArgHisArg-----CysProThrLeuThrGlnAspGluVal 156  
 Db 357 AAATGGCTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 416  
 QY 157 GlnTrpCysAspHisSerSerLeuGlnProSerThrProGluLeuIleLysHisProAla 176  
 Db 417 GAG-----GCGGGAGGAGCGCTTGAGCCCGCAGGAGTT---TGAGTGAGCCATGATTCGCC 467

QY 177 SerAlaSerGlnValAlaGlyThrLysAspMetHisHisTyrrThrTrpLeuLeuPheIle 196  
 Db 468 ACTGAGCGCTGGGTGACAGCCCA-----CCCTGTCTCTCAATTTT 509  
 QY 197 PheIlePheAsnPheLeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrp 216  
 Db 510 TTATTTATTTTATTTTGGAGACACAGTCTCGG-TCTGTGCCCGCAGGCTGGAGTGAGTGG 568  
 QY 217 ArgAsnLeuGlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSer 236  
 Db 569 CGCGATCTCGGCTTACTGCAACCTCCGCTCCCGGTTTCGAGCAGTCTCTCTCGCTCAGC 628  
 QY 237 LeuLeuSerSerTrpAspTyrr---ArgArgProProArgLeuAlaAsnPhePheValPhe 255  
 Db 629 CTCGAGTAGTGTGGATTGGCGGCGCGCCACCATCCGGCTAAAT---TTTGTATTT 685  
 QY 256 LeuValGluMetGly-PheThrMetPheAlaArgLeuLeuLeuSerGlyPro--Cys 274  
 Db 686 TTAGTGGAGACGGGCTTTTTCGTTGGTTCAGGCTGGTCTTGAATTTCCCAACCTCAGGT 745  
 QY 275 AspLeuProAlaSerAlaSerGlnSerAlaGlyIleThrGlyValSerHisAlaArg 294  
 Db 746 GATCCACCGCTCAGCTCCCAAAGTGTGGGATTACAGGCATCAGCCACCGCACC 805  
 QY 295 LeuIlePheAsn-----PheCys-----LeuPhe 302  
 Db 806 CAAAGTTTAAATGAATAAATTTAAAGAGAGTCTAGTTTTTTTGGCAATTTTGT 865  
 QY 303 GluMetGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGlySerLeu 322  
 Db 866 TTGACACTCTGCT-CTTGTTCGCCGGCTGGAGTGCAGTGCACAACTCGGCTCGCTG 924  
 QY 323 GlnProLeuProGlyLeuLysArgPheSerCysLeuSerLeuProSerSerTrpAsp 342  
 Db 925 CAATCTCGCTCCCGGTTCAAGTAATTC---TGCTTAGCTACTCAGTGGCTGGAT 981  
 QY 343 TyrrGlyHisLeuProProHisProAlaAsnPheCysIlePheIleArgGlyValSer 362  
 Db 982 TGCAGGCGCCACCATGCCAGTAACTTTTGTGTGTAGTGGAGACAGGGTTTCA 1041  
 QY 363 ProTyrrLeuSerGlyTrpSerGlnThrProAspLeuArg 375  
 Db 1042 CCGTTCGCGCGGATGTTTGAACCTCTGACTTCAGA 1080

AW600804 990 bp mRNA linear EST 23-MAR-2000  
 N26-20 Human leukocyte MATCHMAKER cDNA Library Homo sapiens cdna,  
 mRNA sequence.  
 ACCESSION AW600804 GI:7304735  
 VERSION AW600804.1  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 990)  
 AUTHORS Meng, X., Song, Y., Mao, J., and Sun, Z.  
 TITLE Proteins that interact with NBS  
 JOURNAL Chin. J. Radiol. Med. Prot. (2000) In press  
 COMMENT Contact: Meng X  
 Department of Biochemistry and Molecular Biology  
 Beijing Institute of Radiation Medicine  
 27 Taiping Road, Beijing, 100850, P.R. China  
 Tel: 861068214653  
 Fax: 861068214653  
 Email: xiangbing@hotmail.com.  
 Location/Qualifiers  
 1..990  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /strain="Caucasian"

FEATURES  
 source

```

/db_xref="taxon:9606"
/sex="male"
/cell_type="peripheral blood leukocytes"
/dev_stage="pooled from 6hr & 24hr PHA stimulated
peripheral blood leukocytes"
/clone_lib="Human leukocyte MATCHMAKER cDNA Library"

ORIGIN
Alignment Scores:
Pred. No.: 1,48e-37 Length: 990
Score: 654.50 Matches: 173
Percent Similarity: 52.62% Conservative: 28
Best Local Similarity: 45.29% Mismatches: 90
Query Match: 32.18% Indels: 92
DB: 2 Gaps: 14

US-09-380-203-2 (1-375) x AW600804 (1-990)
QY 11 GluCysAsnGlyValAlaSerAlaHisArgAsnLeuArgLeuProGlySerSerAspSer 30
Db 988 GAGTACAGTTGGCAACCTTGGCTCAGTCACTCACTCACTCCACCTCTCAGGTTTAAGCAATCT 929
QY 31 ProAlaSerAlaSerProValAlaGlyLeuThrGlyMetCysThrHisAlaArgLeuLeu 50
Db 928 CCTGCTCAGCCTCTCTGAGCAGTGGGATTGCGAGTGTGCACCCACACCCAGTTA--- 872
QY 51 LeuTyrPhePheLeuValGluMetGluPheLeuHisValGlyGlnAlaGlyLeuGluLeu 70
Db 871 AGATTGTATTTTATGATGAGACGGGTTTCCACATGTTGGCGAGATGGTCTTGAATCTC 812
QY 71 ProThrSerAspAspProSerValSerAlaSerGlnSerAlaArgTyrArgThrGlyHis 90
Db 811 CTGACCTCAGGTGATCCACCCGCTCGGCTCCCAAGT----- 773
QY 91 HisAlaArgLeuCysLeuAlaAsnPheCysGlyArgAsnArgValSerLeuMetCysPro 110
Db 772 -----GATGGAATTACAGCGTGAACACCTC----- 746
QY 111 SerTyrSerProGluLeu-----LysGln 118
Db 745 GCCTGG-----CCAACATCAGGATAGTATTTTCATGTCATCTGATAGATAAGTCA 689
QY 119 SerThrCysLeuSerLeuProLysCysTrpAspTyrArgArgAlaAlaValProGlyLeu 138
Db 688 CAGAGGTGCTCATTTTCCTCGAA----- 665
QY 139 PheLeuLeuPheLeuArgHisArg---CysProThrLeuThrGlnAsp----- 154
Db 664 -----TGACACTTGTAAATGTCCAAACTATCGAGTGAGATCCCAAG 623
QY 155 -----GluValGlnTrpCysAspHisSerSerLeuGlnProSerThrProGluLeu 171
Db 622 AGCATCAGTCATCTCCATCGG----- 602
QY 172 LysHisProProAlaSerAlaSerGlnValAlaGlyThrLysAspMetHisHisTyrThr 191
Db 601 ---CATCTCCTCGCAT-----CTAGGCATCTCAATTTATTTATTTATTTATTTATTT 557
QY 192 TrpLeuLeuPheLeu---PheLeuPheAsnPhe----- 201
Db 556 TATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 497
QY 202 -----LeuArgGlnSerLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeu 219
Db 496 TTTATTATGAGACGAGTCTTGC-TCTGTGCGCCAGTCTGGAGTGCAGTGGCGCAGTCTC 438
QY 220 GlySerLeuGlnProLeuProGlyPheLysLeuPheSerCysProSerLeuLeuSer 239
Db 437 TGCTCAGTCAAGACTCTGCTCCAGTTTCAGCCATCTCTCTGCTCAGCCTCTCGAGT 378
QY 240 SerTyrAspTyrArgArgPro---ProArgLeuAlaAsnPheValPheLeuValGlu 258
Db 377 AGCTGGGACTACAGGCGCCACGACCGAACCAGCAGCTAAATTTTGTATTTTATGATGAGA 318

```

```

QY 259 MetGlyPheThrMetPheAlaArgLeuLeuLeuSerGlyProCysAspLeuProAla 278
Db 317 CGGGTTTTCACCATGTTAGCCAGGATGGTTTCGATCTCTGACCTTGATTCGCCGCC 258
QY 279 SerAlaSerGlnSerAlaGlyLeuThrGlyValSerHisHisAlaArgLeuLeuPheAsn 298
Db 257 TCAGCCTCCCAAGTGTCTGGGATTACAGCGTGAGCCACCGTGCCGGCTATTTTATTT 198
QY 299 PheCysLeuPhe-----GluMetGluSerHisSerValThrGlnAlaGlyValGln 315
Db 197 TTTAAATTTTATTTTGTGGAGACAGATCTATCTATCTCCAGCTCGAGTGCAG 138
QY 316 TrpProAsnLeuGlySerLeuGlnProLeuProGlyLeuLysArgPheSerCysLeu 335
Db 137 TGGTGTGATCTCAGCCCACTGCACCTCCGCTCCGGGTTCAAGTGATTTCTCTGCTC 78
QY 336 SerLeuProSerSerTyrAspTyrGlyHisLeuProProHisProAlaAsnPheCysLeu 355
Db 77 AGCTCCCAAGTAGCTGGGATTACAGGTGTCTGCCACCATGCCAGCTAAATTTTGTATT 18
QY 356 PheLeu 357
Db 17 TTTGTC 12

RESULT 15
LOCUS CR627381
DEFINITION Homo sapiens mRNA; cDNA DKFZp762F237 (from clone DKFZp762F237).
ACCESSION CR627381
VERSION CR627381.1 GI:50949857
KEYWORDS HTC
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3775)
AUTHORS Ottenwaelder,B., Obermaier,B., Deutschenbaun,S., Schaijpp,A.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
The German cDNA Consortium
Direct Submission
Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinried/Germany) within the cDNA
sequencing consortium of the German Genome Project.
This clone (DKFZp762F237) is available at the RZPD Deutsches
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp762F237
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES
Location/Qualifiers
1..3775
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp762F237"
/tissue_type="melanoma (MeWo cell line)"
/clone_lib="762 (synonym: hmel2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="hypothetical protein, N-terminus truncated"
1..3775
/gene="DKFZp762F237"
<13..744
/gene="DKFZp762F237"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH10478.1"
/db_xref="GI:50949858"

```

/translation="YGRNVVVEALQEFWMKOSRGADLKNGLVAVVEMVPSNPVVC  
VYTLPGGSCFGSFCPCPKAEARRSAKIALMNSVNEHPSRIRIDEIEKSVSALA  
SNGNREARNNTGIGAFPMLESNKGMLEFQELMTVTFOLLHNSGLKAMRRQC  
SRQEVLAHSHRALDDDIRHQMALDWDVRSBQSVPGALSRELASTERLEARLAGKEL  
RPHKEKDIILVLAAGQLGNMHSNC"

ORIGIN

Alignment Scores:  
Pred. No.: 2,95e-36 Length: 3775  
Score: 646.50 Matches: 174  
Percent Similarity: 52.04% Conservative: 20  
Best Local Similarity: 46.65% Mismatches: 70  
Query Match: 31.78% Indels: 113  
DB: 3 Gaps: 6

US-09-380-203-2 (1-375) x CR627381 (1-3775)

QY	6	LeuLeuProArgLeuGluCysAsnGlyAlaIleSerAlaHisArgAsnLeuArgLeuPro	25
Db	1905	CTTTTGGCCAGGCTGGAGTGCAGTGGCTGATCTCGGCTCACTGCAACCTCTGCCTCCAG	1964
QY	26	GlySerSerAspSerProLaserAlaSerProValAlaGlyIleThrGlyMetCysThr	45
Db	1965	GGTTCGAAGCGATTCTCTTGCCTCAGCTCCGAGTAGCTGGGATTACAGGC---ACACGC	2021
QY	46	HisAlaArgLeuIleLeuTyrrPhePheLeuValGluMetGluPheLeuHisValGlyGln	65
Db	2022	CACAACGGCTGGCTTAATTTATATTTTATAGAGACAGGTTTTCACCATGTTGGCCAG	2081
QY	66	AlaGlyLeuGluLeuProThrSerAspProSerValSerAlaSerGlnSerAlaArg	85
Db	2082	GCTGTTCTTGAAC	2095
QY	86	TyrArgThrGlyHisAlaArgLeuCysLeuAlaAsnPheCysGlyArgAsnArgVal	105
Db	2095	-----	2095
QY	106	SerLeuMetCysProSerTrpSerProGluLeuLysGlnSerThrCysLeuSerLeuPro	125
Db	2096	-----CCTGACATAAGGTAAATCCACCCGCTCAGCCTCCCA	2131
QY	126	LysCysTrpAspTyrArgArgAlaVal--ProGlyLeuPheIleLeuPheLeuA	145
Db	2132	AGTGTGGATTACAGGCTGAGGCACCTGAGCCTGGCCTATTTGGGTTTTTTT---	2187
QY	145	rgHisArgCysProThrLeuThrGlnAspGluValGlnTrpCysAspHisSerSerLeuG	165
Db	2187	-----	2187
QY	165	lnProSerThrProGluIleLysHisProProAlaSerAlaSerGlnValAlaGlyThrL	185
Db	2187	-----	2187
QY	185	ysAspMetHisIstYrThrTrpLeuIlePheIlePheAsnPheLeuArgGlnS	205
Db	2188	-----TGGTT--TTGTTTTTTGTTTTTG-----TTTTTGACACGA	2220
QY	205	erLeuAsnSerValThrGlnAlaGlyValGlnTrpArgAsnLeuGlySerLeuGlnProL	225
Db	2221	GTCTTGC-TCTGTGCGCCAGGCTAGAGTGCAGTGGCGGATCTCGGCTCACTGCAACCTC	2279
QY	225	euProGlyPheLysLeuPheSerCysProSerLeuLeuSerSerTrpAspTyrArg-	244
Db	2280	CGCCTCCCAAGATTCACACATTCTCTGCTCAGCCTCCGAGTAGCTGGGATTACAGGT	2339
QY	245	--ArgProProArgLeuAlaAsnPheValPheLeuValGluMetGlyPheThrMetP	264
Db	2340	GGCCACCACACGCCAGCATATTTTGTATTTTGTAGACGGGGTTTACCATGT	2399
QY	264	heAlaArgLeuIleLeuSerGlyProCysAspLeuProLaserAlaSerGlnSerA	284
Db	2400	TAGCCAGGACGGTCTCCATCTCTGACCTCGTGATCCACCTGCTCGGCTCCCAACTA	2459

Search completed: September 16, 2005, 01:10:29  
Job time : 4052 secs

QY	284	laGlyIleThrGlyValSerHisAlaArgLeuIlePheAsnPheCysLeuPheGluM	304
Db	2460	TT-----TGTTGTTTTT-TTTTTTTTTTTGAGA	2487
QY	304	etGluSerHisSerValThrGlnAlaGlyValGlnTrpProAsnLeuGlySerLeuGlnP	324
Db	2488	CGGAGTCTCGCTCTCTGCCCCAGGCTGGAGTGCAGTGGCGCTATCTCGGCTCACTGCAAG	2547
QY	324	roLeuProProGlyLeuLysArgPheSerCysLeuSerLeuProSerSerTrpAspTyrG	344
Db	2548	CTCCGCTCCCGGGTTCACGCCACTCTCTGCTCAGGCTCCTTAGTAGCTGGGACTACA	2607
QY	344	lyHisLeuProProHisProAlaAsn--PheCysIlePheIleArgGlyGlyValSerPr	363
Db	2608	GGCGCCCGCCACCATGCTGCTAAATTTTTTGTATTTTGTAGTAGACGGGGTTTCACC	2667
QY	363	oTyrLeuSerGlyTrpSerGlnThrProAsp	373
Db	2668	GTGTTTGACAGGATGGTCTCGATTTTCCTGAC	2698